



Méthode ARPEGES (Analyse de Risque Pesticides pour la Gestion des Eaux de Surface). Guide d'aide à l'interprétation des résultats du modèle pour l'Etat des Lieux 2019 de la Directive Cadre sur l'Eau

E. Adoir, Nadia Carluer, Véronique Gouy

► To cite this version:

E. Adoir, Nadia Carluer, Véronique Gouy. Méthode ARPEGES (Analyse de Risque Pesticides pour la Gestion des Eaux de Surface). Guide d'aide à l'interprétation des résultats du modèle pour l'Etat des Lieux 2019 de la Directive Cadre sur l'Eau. [Rapport de recherche] irstea. 2018, pp.143. hal-02607533

HAL Id: hal-02607533

<https://hal.inrae.fr/hal-02607533>

Submitted on 16 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Méthode ARPEGES (Analyse de Risque Pesticides pour la Gestion des Eaux de Surface)

**Guide d'aide à l'interprétation des
résultats du modèle pour l'Etat des Lieux
2019 de la Directive Cadre sur l'Eau**

Rapport final

Emilie ADOIR (Irstea)

Nadia CARLUER (Irstea)

Véronique GOUY (Irstea)

Février 2018

- **AUTEURS**

Emilie ADOIR, ingénieur (Irstea, équipe Pollutions diffuses agricoles)

Nadia CARLUER, ICPEF (Irstea, équipe Pollutions diffuses agricoles)

Véronique GOUY, IDAE (Irstea, équipe Pollutions diffuses agricoles)

- **CONTRIBUTEURS**

Katell MELLAC, ingénieur (Irstea, équipe Laboratoire d'Hydro-écologie Quantitative)

Jérémy PIFFADY, IPEF (Irstea, équipe Laboratoire d'Hydro-écologie Quantitative)

- **CORRESPONDANTS**

Ministère de la Transition écologique et solidaire : Jeanne BOUGHABA, IPEF

Agence française pour la biodiversité : Nolwenn BOUGON, Chargée de mission Pressions/Impacts Chimie

Droits d'usage : accès libre

Niveau géographique : national

Couverture géographique : France

Niveau de lecture : professionnels, experts

- **METHODE ARPEGES (ANALYSE DE RISQUE PESTICIDES POUR LA GESTION DES EAUX DE SURFACE) : GUIDE D'AIDE A L'INTERPRETATION DES RESULTATS DU MODELE POUR L'ETAT DES LIEUX 2019 DE LA DIRECTIVE CADRE SUR L'EAU (EMILIE ADOIR, NADIA CARLUER ET VERONIQUE GOUY)**

- **SYNTHESE POUR L'ACTION OPERATIONNELLE**

Le devenir des produits phytosanitaires dans l'environnement, et notamment leur transfert vers le milieu aquatique, est déterminé par un ensemble de processus complexes, et qui interagissent de façon non linéaire entre eux. En effet, les molécules phytosanitaires ont la particularité d'interagir avec le milieu, et en particulier de pouvoir être dégradées, ou adsorbées, c'est-à-dire retenues sur les particules de sol (notamment matière organique ou argile) après avoir été appliquées et avant de rejoindre le milieu aquatique par transfert hydrique.

Le devenir d'une substance va donc dépendre à la fois :

- Des caractéristiques physiques du milieu (climat, sol, géologie, relief) qui vont conditionner les écoulements dominants susceptibles d'entraîner les molécules vers le milieu aquatique. On conçoit en effet que des produits entraînés par ruissellement auront moins d'opportunité d'être dissipés (dégradés ou adsorbés) que des produits transitant lentement dans le sol.
- Des caractéristiques propres de la molécule. On caractérise notamment le comportement d'une molécule phytosanitaire dans l'environnement par sa facilité à être dégradée (DT50, ou demie vie, exprimée en jours ; d'autant plus forte que la molécule se dégrade lentement) et à être adsorbée (Koc ou coefficient de partage sol/eau, exprimée en mL/g ; d'autant plus fort que la molécule s'adsorbe facilement, et est donc moins mobile).
- De la façon dont la molécule est utilisée : période, doses et lieux d'application.

ARPEGES (Analyse de Risque Pesticides pour la Gestion des Eaux de Surface) est basée sur ce constat, et rend compte des connaissances expertes (réseau bayésien) mobilisables sur le devenir des produits phytosanitaires dans l'environnement et leur transfert hydrique vers les eaux de surface, en fonction de ces trois ensembles de déterminants. ARPEGES permet ainsi de **croiser la vulnérabilité du milieu** au transfert des produits phytosanitaires et la **pression d'usage exercée par ces substances** sur le territoire français métropolitain.

Cette méthode est basée sur la construction d'un **modèle de connaissance des processus de transfert**. Elle a été élaborée pour fournir un potentiel de contamination du milieu pour l'ensemble des substances actives utilisées, à partir des différentes variables influençant ces phénomènes. Elle distingue les processus se produisant en nappe haute (réserves en eau des sols remplies ; qui correspond à la période automne-hivers de la version ARPEGES de 2012) et en nappe basse (réserves en eau des sols plus faibles ; qui correspond à la période printemps-été de la version ARPEGES de 2012). Les processus en jeu sont complexes et les **données imparfaites (types de données hétérogènes, lacunes) à l'échelle nationale**. Afin de fournir une indication du niveau de confiance associé au résultat, le **croisement** des données se fait de façon **probabiliste** à l'aide d'un **système bayésien**. Au total, 18 variables sont agrégées pour obtenir le potentiel de contamination. L'échelle d'étude est le **Bassin Versant de Masse d'Eau (BVME)** défini dans le cadre de la DCE (environ 10 000 BVME sur le territoire métropolitain). **On rappelle que le terme de contamination est ici utilisé dans le sens d'exposition du milieu, sans préjuger des risques écotoxicologiques associés.**

L'application de cette méthode permet donc de cibler les zones potentiellement les plus à risque de contamination des masses d'eau de surface, de connaître les causes potentiellement majoritaires dans la contamination observée et de proposer une première orientation générale de types d'actions pour réduire ce phénomène, selon les voies prépondérantes du transfert.

La méthode, initialement conçue pour un mélange de substances actives (Gauroy et al, 2012), a été adaptée à l'étude d'une seule molécule. On a en particulier recherché une donnée plus précise et spatialisée pour la pression d'utilisation phytosanitaire que celle utilisée initialement pour élaborer la méthode en 2012. Par ailleurs, l'approche par substance permet la confrontation avec les données de surveillance présentée en partie 5 de ce guide, et ainsi de tester la cohérence des résultats des cartes de potentiel de contamination d'ARPEGES avec les cartes de contamination observées. Il faut cependant noter que la méthode ARPEGES ne s'applique qu'aux substances actives, et non à leurs métabolites.

On rappelle que **la maille de la méthode d'interprétation des données de surveillance est l'Hydro-Eco-Région (HER) modifiée, alors qu'il s'agit du bassin versant de la masse d'eau dans le cas de la méthode ARPEGES.** Ces éléments seront importants à prendre en considération lors de la confrontation des résultats.

Pour aider les Agences de l'Eau à prendre en main le modèle et à interpréter ses résultats, une application sur 15 substances actives prioritaires au regard de la DCE a été réalisée. Des clés d'interprétation s'appuyant sur des exemples obtenus à partir de ces 15 substances sont présentées à la fin de ce guide. En complément de ce travail « thématique », un travail de programmation visant à transcrire le réseau bayésien d'ARPEGES dans un outil libre (R) et à produire une interface simple d'utilisation a été réalisé et rapporté dans un guide utilisateur (Mellac et Piffady, 2018).

Outre la nouvelle source de donnée d'entrée pour la pression phytosanitaire adaptée à l'étude d'une seule substance, d'autres modifications ont été apportées à la méthode produite et appliquée en 2012. Les changements sont présentés au fur et à mesure de ce guide dans les encadrés verts. Ils ont porté à la fois :

- sur la structure graphique du réseau bayésien : suppression de plusieurs variables (aléa érosif, solubilité des substances, densité de réseau hydrographique et vulnérabilité intrinsèque à la dérive) et augmentation de modalités pour la pression d'usage phytosanitaire et les potentiels de contamination ;
- sur la structure probabiliste du réseau : construction de nouvelles tables de probabilités conditionnelles suite à l'évolution de la structure graphique ;
- sur les données d'entrée : pression d'usage phytosanitaire.

Quelques éléments de vocabulaire ont également été modifiés. En voici la correspondance :

Nouveau vocabulaire	Ancien vocabulaire
Nappe haute	Automne-hiver
Nappe basse	Printemps-été
Transferts lents	Chronique
Transferts rapides	Aigu
Potentiel de contamination	Risque de contamination

SOMMAIRE

METHODE ARPEGES (ANALYSE DE RISQUE PESTICIDES POUR LA GESTION DES EAUX DE SURFACE) : GUIDE D'AIDE A L'INTERPRETATION DES RESULTATS DU MODELE POUR L'ETAT DES LIEUX 2019 DE LA DIRECTIVE CADRE SUR L'EAU 9

1. Objectifs de ce guide 9

2. Substances actives étudiées 10

2.1. Critères de choix des substances 10

2.2. Description des caractéristiques physico-chimiques et usages à l'échelle nationale..... 11

2.3. Description des données du SDeS et points de vigilance pour l'analyse.... 11

3. La méthode ARPEGES : description du modèle, de ses entrées et types de sorties..... 15

3.1. Finalité de la méthode : calcul de potentiels de contamination des eaux de surface pour hiérarchiser les causes de la contamination..... 15

3.2. Points saillants de la méthode / La méthode dans ses grandes lignes 16

3.2.1. Echelle de calcul et implications 16

3.2.2. Prise en compte de la vulnérabilité et de la pression..... 17

3.2.2.1. La méthode ARPEGES : une adaptation de la méthode Corpen..... 17

3.2.2.2. Décomposition en vulnérabilités intrinsèques, vulnérabilités spécifiques et potentiels de contamination..... 17

3.2.3. Résultat final : obtention de 4 types de potentiels de contamination 19

3.2.3.1. Notion de saison dans ARPEGES : 19

3.2.3.2. Distinction des vulnérabilités spécifiques en deux types de contamination ... 20

3.2.3.3. Synthèse des différentes variables calculées par ARPEGES..... 21

3.2.4. Approche probabiliste du modèle 23

3.3. Construction du réseau bayésien d'ARPEGES..... 24

3.3.1. Construction de la structure graphique du réseau..... 24

3.3.1.1. 1^{ère} étape : Choix des variables d'intérêt..... 24

3.3.1.2. 2^{ème} étape : Choix des états pris par les variables d'intérêt 31

3.3.1.3. 3^{ème} étape : Etablissement des liens entre variables 31

3.3.2. Construction de la structure probabiliste..... 43

3.4. Données d'entrée 48

3.4.1. Préparation des données d'entrée 48

3.4.1.1. Synthèse des données d'entrée..... 48

3.4.1.2. Collecte des données d'entrée : description et limites 48

3.4.1.3. Ajustement à la maille de calcul d'ARPEGES..... 60

3.4.2. Intégration des données d'entrée : discrétisation des données d'entrée et traduction en n-uplets de probabilités 63

3.5. Types de sorties du réseau bayésien d'ARPEGES..... 64

3.5.1. Sorties brutes 65

3.5.1.1. Fichiers de sortie par saison 65

3.5.1.2. Lecture des n-uplets de probabilités et calcul opérés pour obtenir ces n-uplets..... 69

3.5.1.3. Interprétation des n-uplets : probabilité majoritaire et indice de confiance..... 70

1.1.1.1. 72

1.1.1.2. 72

1.1.1.3. 72

3.5.2. Présentation des résultats sous forme de cartes 73

3.5.3. Sorties communes à toutes les substances actives 75

3.6. Application de la méthode via l'interface R..... 80

4. Méthode de comparaison des sorties ARPEGES à des données observées	81
4.1. Traitement des données observées comparées : méthode d'intégration des courbes enveloppes (sur 3 ans)	81
4.2. Sorties ARPEGES comparées.....	84
5. Test de cohérence du modèle	84
5.1. Résultats de la comparaison et analyse.....	84
5.2. Cas particuliers incohérents	88
6. Etude de cas et pistes d'interprétations	88
6.1. Etude du potentiel de contamination.....	89
6.2. Causes du potentiel de contamination.....	90
6.3. Causes de la vulnérabilité spécifique.....	94
6.4. Liens entre potentiels de contamination via transferts lents ou rapides.....	95
7. Conclusion.....	96
1.1. La méthode ARPEGES et son application au niveau national	96
1.2. L'interface ARPEGES et ses fonctionnalités pour une application locale	97
1.3. Vers une agrégation des résultats pour faciliter une interprétation plus globale ?	98
Bibliographie.....	100
8. Table des illustrations	101
9. Annexes	103
9.1. Annexe 1 : Méthode d'établissement des coefficients de saisonnalité	103
9.1.1. Rappel définition	103
9.1.2. Précisions sur les termes « nappe haute » et « nappe basse »	103
9.1.3. Méthode de détermination des coefficients de saisonnalité	103
9.1.4. Tableau de synthèse des coefficients de saisonnalité déterminé selon cette méthode.....	105
9.1.5. Détermination des coefficients de saisonnalité pour chaque substance active	106
9.2. Annexe 2 : Fiche synthétique pour estimer la variable « Ruissellement/Infiltration »	110
9.2.1. Objectifs	110
9.2.2. Données sources.....	110
9.2.3. Calcul à l'échelle du BVME.....	111
9.2.4. Rappel des modalités que peut prendre la variable dans ARPEGES ...	111
9.2.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES	111
9.3. Annexe 3 : Fiche synthétique pour estimer la variable « Réserve utile » ...	113
9.3.1. Objectifs	113
9.3.2. Données sources.....	113
9.3.3. Calcul à l'échelle du BVME.....	114
9.3.4. Rappel des modalités que peut prendre la variable dans ARPEGES ...	114
9.3.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES	114
9.4. Annexe 4 : Fiche synthétique pour estimer la variable « drainage »	116
9.4.1. Objectifs	116
9.4.2. Données sources.....	116
9.4.3. Calcul à l'échelle du BVME.....	117
9.4.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES	117
9.5. Annexe 5 : Fiche synthétique pour estimer la variable « Hydromorphie »	119

9.5.1.	Objectifs	119
9.5.2.	Données sources.....	119
9.5.3.	Calcul à l'échelle du BVME.....	120
9.5.4.	Rappel des modalités que peut prendre la variable dans ARPEGES ...	120
9.5.1.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	120
9.6.	Annexe 6 : Fiche synthétique pour estimer la variable « Battance »	122
9.6.1.	Objectifs	122
9.6.2.	Données sources.....	122
9.6.3.	Calcul à l'échelle du BVME.....	124
9.6.4.	Rappel des modalités que peut prendre la variable dans ARPEGES ...	124
9.6.1.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	124
9.7.	Annexe 7 : Fiche synthétique pour estimer la variable« Ripisylves 30 m » 126	
9.7.1.	Objectifs	126
9.7.2.	Données sources.....	126
9.7.3.	Calcul à l'échelle du BVME.....	126
9.7.4.	Rappel des modalités que peut prendre la variable dans ARPEGES ...	126
9.7.5.	Résultats et limites.....	126
9.8.	Annexe 8 : Fiche synthétique pour estimer la variable « Ripisylves 10 m » 128	
9.8.1.	Objectifs	128
9.8.2.	Données sources.....	128
9.8.3.	Calcul à l'échelle du BVME.....	128
9.8.4.	Rappel des modalités que peut prendre la variable dans ARPEGES ...	128
9.8.1.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	128
9.9.	Annexe 9 : Fiche synthétique pour estimer la variable « bande enherbée » 130	
9.9.1.	Objectifs	130
9.9.2.	Données sources.....	130
9.9.3.	Calcul à l'échelle du BVME.....	130
9.9.4.	Rappel des modalités que peut prendre la variable dans ARPEGES ...	130
9.9.1.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	130
9.10.	Annexe 10 : Fiche synthétique pour estimer la variable «Zone climatique» 132	
9.10.1.	Objectifs	132
9.10.2.	Données sources	132
9.10.3.	Calcul à l'échelle du BVME.....	132
9.10.4.	Rappel des modalités que peut prendre la variable dans ARPEGES .	132
9.10.1.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	132
9.11.	Annexe 11 : Fiche synthétique pour estimer la variable « Cumul de pluies » 134	
9.11.1.	Objectifs	134
9.11.2.	Données sources	134
9.11.3.	Calcul à l'échelle du BVME.....	134
9.11.4.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	135
9.12.	Annexe 12 : Fiche synthétique pour estimer la variable « Teneur en matière organique du sol »	137
9.12.1.	Objectifs	137
9.12.2.	Données sources	137
9.12.3.	Calcul à l'échelle du BVME.....	137
9.12.4.	Rappel des modalités que peut prendre la variable dans ARPEGES .	138
9.12.1.	Résultat de la discrétisation – Carte en entrée d'ARPEGES	138
9.13.	Annexe 13 : Fiche synthétique pour estimer la variable « Pression d'usage phytosanitaire »	139
9.13.1.	Objectifs	139
9.13.2.	Données sources	139
9.13.3.	Calcul à l'échelle du BVME.....	139
9.13.4.	Rappel des modalités que peut prendre la variable dans ARPEGES .	139

9.13.5. Résultat de la discrétisation – Carte en entrée d'ARPEGES	139
9.14. Annexe 14 : Vulnérabilités spécifiques pour les 9 combinaisons de caractéristiques physico-chimiques des substances.....	140
9.14.1. Période de Nappe Haute. Transferts rapides	140
9.14.2. Période de Nappe Haute. Transferts lents	141
9.14.3. Période de Nappe Basse. Transferts rapides	142
9.14.4. Période de Nappe Basse. Transferts lents	143

METHODE ARPEGES (ANALYSE DE RISQUE PESTICIDES POUR LA GESTION DES EAUX DE SURFACE) : GUIDE D'AIDE A L'INTERPRETATION DES RESULTATS DU MODELE POUR L'ETAT DES LIEUX 2019 DE LA DIRECTIVE CADRE SUR L'EAU

1. Objectifs de ce guide

Ce guide a trois objectifs principaux :

- Mettre à jour et rendre plus transparente la description de la méthode ARPEGES, afin de rendre l'utilisateur plus autonome.
- De tester la cohérence des sorties d'ARPEGES à des données observées.
- Donner des clés d'interprétation des sorties d'ARPEGES à travers des exemples pour des substances actives d'intérêt pour l'Etat Des Lieux 2019.

Il s'agit d'un guide dont la lecture n'est pas obligatoirement linéaire, mais dépend de ce que recherche le lecteur :

- **Partie 4.1 et 4.2** : « Je comprends la méthode dans ses grandes lignes »
- **Partie 4.3** : « Je comprends le moteur¹ d'ARPEGES »
- **Partie 4.4** : « Je prends connaissance des données d'entrée utilisées dans le projet ARPEGES-ELD 2019 »
- **Partie 4.5** : « Je comprends le calcul des sorties »
- **Partie 5 et 6** : « J'évalue la cohérence des sorties aux données observées »
- **Partie 7** : « J'analyse les sorties dans un objectif de hiérarchiser les causes de la contamination »

Tout au long de la partie 4 (méthode), des encadrés thématiques sont par ailleurs ajoutés pour repérer certains types d'information :

- **Les fondements mathématiques** : ces encadrés font figurer à part les explications théoriques des réseaux bayésiens, présentées pas-à-pas et sur des exemples simples.
- **Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 »** : ces encadrés pointent les éléments modifiés, et rappellent de quoi était composée la méthode auparavant.
- **Les possibilités de l'interface R** : ces encadrés font des liens avec le guide utilisateur dédié (Mellac et Piffady, 2018).

NB : On alerte le lecteur dès à présent sur une question de vocabulaire : le terme « **pression** » n'est employé dans ce guide que pour signifier la **pression liée aux usages phytosanitaires**, c'est-à-dire la **quantité de substance active appliquée par unité de surface (il s'agit donc d'une donnée en g/ha)**. Il s'agit d'une pression potentielle, avant transfert vers le cours d'eau. Il ne s'agit en aucun cas de la pression exercée sur les masses d'eau, qui constitue le résultat du transfert des substances actives vers les cours d'eau.

¹ On entend par moteur la « mécanique interne » d'ARPEGES, ie la façon dont les données d'entrée sont croisées pour conduire au résultat final.

2. Substances actives étudiées

2.1. Critères de choix des substances

Quinze substances actives ont été étudiées dans la mise en application d'ARPEGES pour l'EDL 2019 (Tableau 1). La sélection de ces substances s'appuie sur les besoins des bassins, discutés lors de groupes de travail nationaux sur les pollutions diffuses agricoles dans le cadre de la mise en œuvre de la DCE et SDAGE, animés par le ministère en charge de l'écologie.

Elle s'appuie sur 2 critères, appliqués aux substances prioritaires DCE (exception faite du S-métolachlore qui n'est pas prioritaire DCE):

- La substance est considérée d'intérêt par au moins une Agence

OU

- La substance a un taux de quantification significatif et parmi les plus élevés à l'échelle nationale en se basant sur les données de l'année 2014 (données disponibles les plus récentes).

Substances actives	Taux de quantification en 2014 (France entière)	Demande AESN	Demande AELB	Demande AEAG	Demande AERM
Glyphosate	39%		X		X
Diflufénicanil	18%		X		X
Chlortoluron	17%		X		X
Bentazone	17%	X		X	
S-metolachlore (non prio DCE)	16%		X		X
Boscalid	15%				
Métaldéhyde	14%		X	X	
Métazachlore	13%		X		X
Imidaclopride	11%				X
Aminotriazole (Amitrole)	9%			X	X
2,4-mcpa	9%	X		X	X
2,4-d	9%	X			X
Oxadiazon	7%			X	

Nicosulfuron	6%	X	X	X	X
Pendiméthaline	2%		X		

Tableau 1 : Liste des substances actives retenues pour l'étude et classement selon les critères de sélection

L'isoproturon a été retiré de la liste car interdit depuis 2016. Cependant, le GT SDAGE a souhaité maintenir l'aminotriazole pour l'étude, malgré son interdiction depuis 2015.

Bien que non prioritaire DCE, le S-métolachlore fait partie des substances les plus quantifiées en France, d'où le maintien dans la liste.

A noter que trois substances n'ont, a posteriori, présenté que peu d'intérêt dans ce projet, car des difficultés inhérentes à la donnée d'entrée « pression phytosanitaire » ont empêché d'apprécier de manière satisfaisante les quantités appliquées sur les bassins versants de masse d'eau (notion de « quantités orphelines », cf 3.4.1.2). Il s'agit de l'oxadiazon, du métaldéhyde et de l'imidaclopride. Cette information est donc importante à prendre en compte avant de débiter une évaluation avec ARPEGES.

2.2. Description des caractéristiques physico-chimiques et usages à l'échelle nationale

Les informations sont rassemblées dans le Tableau 2.

Les caractéristiques physico-chimiques sont extraites de la base de données SIRIS, qui est la source utilisée pour le K_{oc} et la DT_{50} dans ARPEGES (cf 3.4.1).

Les principaux usages agricoles ont été relevés dans la base de données MatPhyto, qui est une compilation du guide ACTA, et dont l'entrée de recherche est la substance active (à la différence de l'outil E-phy de l'Anses dont l'entrée est le produit phytosanitaire).

Enfin, les quantités globales vendues en France en 2014 et 2015 ont été obtenues à partir de la Base Nationale des Ventes des Ventes de produits phytopharmaceutiques par les Distributeurs agréés (BNV-d).

2.3. Description des données du SDeS et points de vigilance pour l'analyse.

Les données exploitées pour évaluer la cohérence des sorties d'ARPEGES avec des données observées (cf partie 5) sont des données de suivi de la qualité des eaux acquises par les Agences de l'Eau, et centralisées par le SDeS, service de statistiques du Ministère de l'Environnement. Les données étudiées couvrent la période 2013-2015, afin de lisser la variabilité interannuelle des concentrations en substances actives dans les cours d'eau.

Ce paragraphe a pour but de fournir des informations très synthétiques sur le contenu de cette base de données.

La **Erreur ! Source du renvoi introuvable.** confirme les chiffres indiqués dans le Tableau 1 pour le t aux de quantification des substances actives à l'échelle nationale, mais pour la période 2013-2015. Le glyphosate présente même sur la période un taux de quantification supérieur à 65%.

La Figure 2 donne une information spatialisée sur la pression de mesure : les quinzaines où plus de 10 mesures ont été effectuées pour chaque substance (parmi les 15 étudiées ici) ont été dénombrées par Hydro-Eco-Région (HER) sur la période 2013-2015, puis une moyenne et un écart-type ont été calculés par HER pour les quinze substances. Cette information est utile pour l'étude des 90èmes centiles moyens des concentrations (cf 4.1) : on peut ainsi connaître quelles sont les HER où le calcul du 90^{ème} centile a plus de sens (i.e. lorsque le nombre de mesure par quinzaine est supérieur à 10).

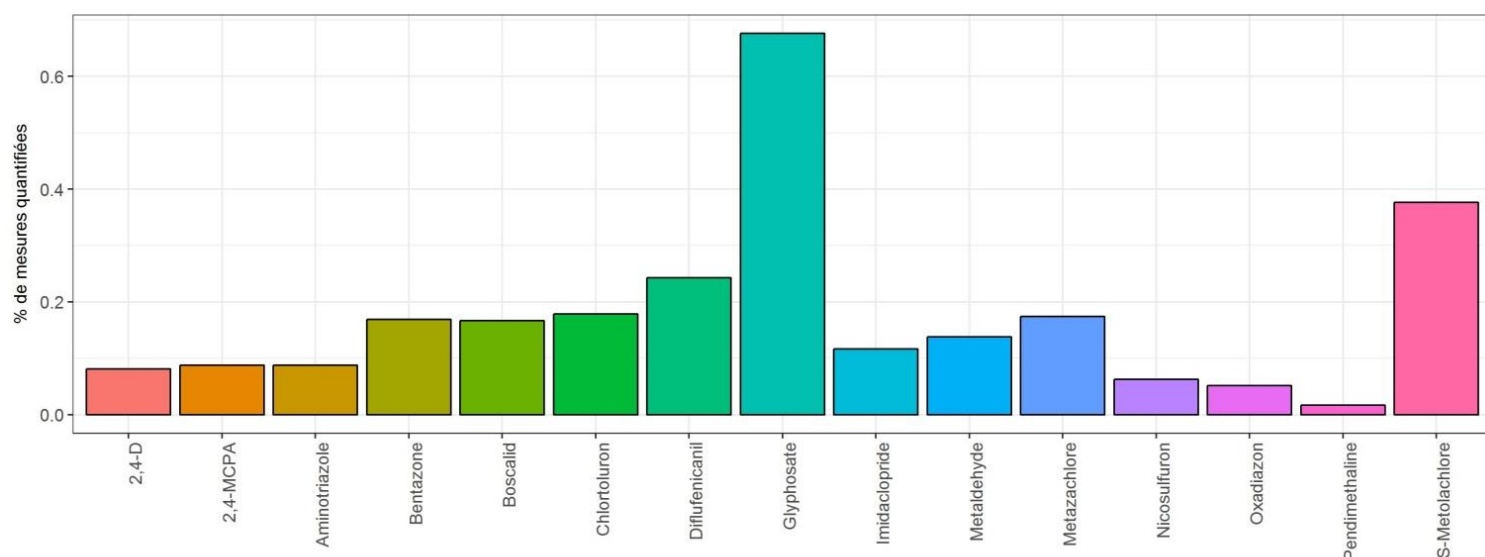


Figure 1 : Proportion de quantification des substances actives étudiées à l'échelle nationale sur la période 2013-2015

Substances actives	Groupe	Koc ² (mL.g-1)	DT50 champ ² (jours)	Solubilité ² (mg.L-1)	Usages agricoles principaux ³	Autres usages ³	Quantités totales vendues en France en 2015 ⁴ (kg)	Ecart de vente 2015-2014 ⁴ (kg)
Glyphosate	herbicide	21699,44	31,5	10500	toutes cultures		8 452 097	-1 019 753
Diflufénicanil	herbicide	3417	315	0,05	blé tendre, orge, autres céréales, fourrages (ray grass), divers (lavande)	cultures ornementales	413 515	89 744
Chlortoluron	herbicide	208	34	74	blé tendre, orge, autres céréales, légumes-fleurs (pavot, œillets), divers (lupin porte-graine)		1 275 260	129 270
Bentazone	herbicide	51,5	10	570	blé tendre, orge, autres céréales, maïs, autres oléagineux, protéagineux, fourrages, légumes-fleurs, divers (lin textile, PPAM)	cultures ornementales	202 984	-8 083
S-metolachlore (non prioritaire DCE)	herbicide	226	21	480	cultures fourragères (betterave), autres cultures industrielles (betterave), maïs, tournesol, autres céréales (sorgho), autres oléagineux (soja), légumes-fleurs		2 024 737	-12 499
Boscalid	fongicide	809	200	4,64	blé tendre, orge, autres céréales, colza, autres oléagineux, vergers, légumes-fleurs, vignes, divers (porte- graines)		277 171	-87 191
Métaldéhyde	molluscicide	85	4,4	200		traitement de sols (anti- limaces)	750 682	-84 010
Métazachlore	herbicide	110	9,8	630	colza, tournesol, légumes-fleurs, divers (plantes à parfum PPAM)	cultures ornementales	707 081	4 180
Imidaclopride	insecticide	225	174	610	vergers	traitement de semences, forêts, traitement de sols, cultures ornementales	261 744	4 375

² Source : SIRIS, 2012 (https://siris-pesticides.ineris.fr/bdd_siris_pesticides)

³ Source : MatPhyto, 2017 (<http://index-matphyto.univ-lyon1.fr/>)

⁴ Source : BNV-d code postal distributeur, 2014 et 2015

Aminotriazole	herbicide	91	18	264000	maïs, vergers, vignes, traitements généraux		52 407	-31 682
2,4-mcpa	herbicide	74	25	293900	blé tendre, orge, autres céréales, prairies permanentes, riz, fourrage	cultures ornementales	673 940	-62 741
2,4-d	herbicide	56	9,9	23180	blé tendre, orge, autres céréales, vergers, fruits à coques, légumes-fleurs (asperges), riz, prairies permanentes, prairies temporaires, divers (PPAM), zones cultivées après récolte	forêts, cultures ornementales	544 290	-83 738
Oxadiazon	herbicide	2339	105	1	tournesol, autres oléagineux (soja), vergers, légumes-fleurs, riz, divers (PPAM)	cultures ornementales	2 711	-4 092
Nicosulfuron	herbicide	20,7	19,3	7500	maïs		51 198	-3 138
Pendiméthaline	herbicide	15744	99,17	0,33	blé tendre, orge, autres céréales, maïs, tournesol, autres oléagineux, protéagineux, vergers, légumes-fleurs, autres cultures industrielles (tabac), divers		1 063 974	135 232

Tableau 2 : Description des caractéristiques physico-chimiques des substances étudiées et de leurs usages

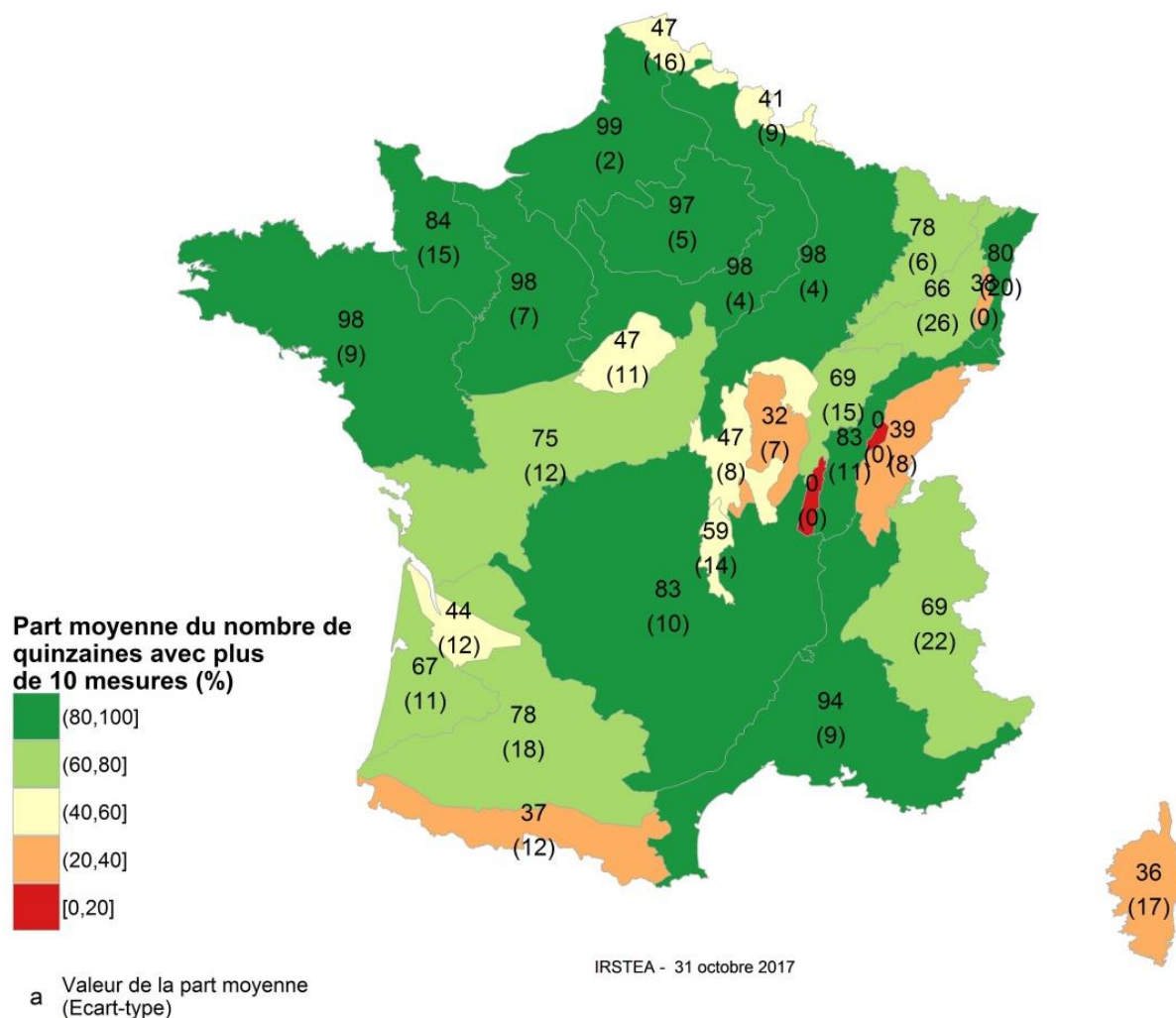


Figure 2 : Pourcentage du nombre de quinzaines de jours où plus de 10 mesures ont été effectuées par HER (moyenne sur la période 2013-2015 et pour les quinze substances actives étudiées)

3. La méthode ARPEGES : description du modèle, de ses entrées et types de sorties

3.1.Finalité de la méthode : calcul de potentiels de contamination des eaux de surface pour hiérarchiser les causes de la contamination

La démarche proposée est basée sur une approche descriptive et experte des bassins versants dans leur fonctionnement global. Elle repose sur une hiérarchisation des causes d'altération de l'état des cours d'eau liées à la pression exercée par une substance active donnée, en considérant les aspects essentiels de la contamination : la vulnérabilité du milieu aux transferts, les usages (ou pression phytosanitaire), et les propriétés de la substance considérée.

Chacun de ces aspects est décrit à partir d'informations spatialisées, structurées et croisées à l'aide des méthodes bayésiennes. Ces informations, agrégées aux échelles d'intérêt à partir de techniques

SIG, sont issues des données disponibles sur l'ensemble du territoire métropolitain pour assurer l'homogénéité de la méthode. La modélisation bayésienne, probabiliste, repose sur l'expression de l'état d'une ou plusieurs variables en fonction de l'état d'autres variables, dites explicatives, via l'expression de probabilités conditionnelles. Par exemple : « Sachant que les sols sont très battants sur un bassin versant et que les bords de rivière ne sont pas protégés par une ripisylve ou des zones tampons enherbées, alors il y a de fortes chances que le ruissellement hortonien advienne fréquemment et atteigne la rivière ». La probabilité d'occurrence du ruissellement est dans ce cas de figure considérée comme plus élevée.

Cette structure de modèle a l'avantage d'être transparente, et permet de remonter plus facilement aux causes de la contamination qu'un modèle mécaniste, avec les limites méthodologiques que cette structure comporte. Notamment, les données d'entrée doivent être discrétisées, ce qui implique des effets de seuils de classes, et le découpage temporel simplifié rend la prise en compte du décalage temporel entre application et transfert plus difficile.

On précise ici que la méthode ARPEGES étudie uniquement les substances actives, et n'est pas adapté pour l'évaluation du potentiel de contamination des métabolites de ces substances. En effet, les processus intervenant dans la contamination par les métabolites sont encore trop peu connus pour pouvoir être intégrés dans le modèle. Il faut donc prendre des précautions dans les conclusions prises suite à l'étude des résultats d'ARPEGES : des substances peuvent apparaître avec un faible potentiel de contamination dans ARPEGES, mais leurs métabolites peuvent présenter un potentiel de contamination important.

3.2. Points saillants de la méthode / La méthode dans ses grandes lignes

3.2.1. Echelle de calcul et implications

L'échelle de restitution de la méthode ARPEGES est celle des bassins versants locaux de masse d'eau ou BVME (Bassin Versant de Masse d'Eau).

On définit le bassin versant local de la masse d'eau (la « masse d'eau » étant l'échelle de restitution de l'Etat des lieux DCE) comme le bassin versant de son exutoire, amputé le cas échéant des bassins versants des masses d'eau amont (Figure 3). Pour un bassin versant d'une masse d'eau amont (affluent de cours d'eau plus importants), le BVME correspond au bassin versant entier.

En effet, la méthode ARPEGES a pour objectif d'évaluer le potentiel relatif de contamination d'une masse d'eau de surface suite aux applications réalisées sur son bassin versant local : la contribution des masses d'eau amont n'est ainsi pas prise en compte dans le potentiel de contamination du BV local.

L'avantage de fournir une note de potentiel de contamination par BVME est double, il permet de :

- s'affranchir des problèmes d'estimation de la dilution des concentrations provenant des masses d'eau amont
- cibler les zones où instaurer en priorité des plans d'actions.

En contrepartie, la confrontation avec les données de surveillance pour les stations situées sur les grands cours d'eau peut ne pas être pertinente, car un tronçon d'un tel cours d'eau est alimenté par son bassin versant local et des bassins versants amont.

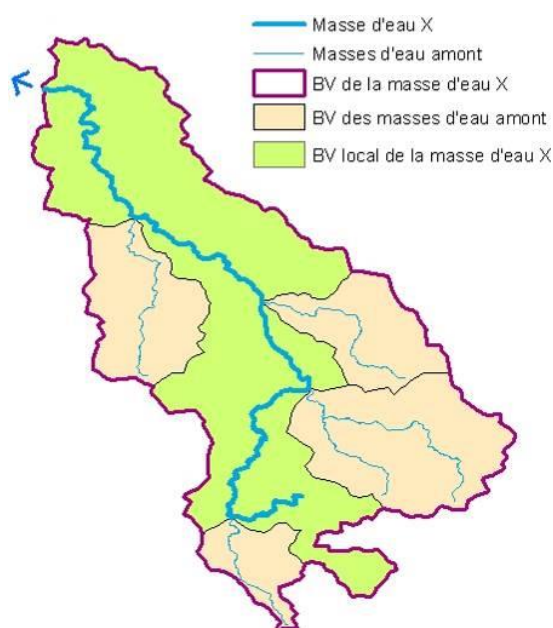


Figure 3: Illustration de la notion de bassin versant de la masse d'eau (BVME). Données : Agence de l'eau Loire-Bretagne

3.2.2. Prise en compte de la vulnérabilité et de la pression

3.2.2.1. La méthode ARPEGES : une adaptation de la méthode Corpen

La méthode ARPEGES s'appuie sur la démarche générale de diagnostic proposée en 2003 par le Comité d'orientation pour des pratiques agricoles respectueuses de l'environnement (Corpen). Celui-ci définit le potentiel de contamination comme résultant d'une **conjonction** entre la **vulnérabilité des masses d'eau** et **l'utilisation de produits phytosanitaires** (« pression phytosanitaire »).

Toutefois, alors que la méthode Corpen consistait surtout à offrir un « cadre » de réflexion aux experts locaux à même d'identifier les processus dominants sur la région considérée, de définir les croisements de données permettant effectivement de mettre en exergue les situations discriminantes, et de choisir les seuils définissant les différentes classes de données de façon pertinente, il a ici fallu effectuer ce travail en amont, afin d'aboutir à une démarche harmonisée au niveau national.

3.2.2.2. Décomposition en vulnérabilités intrinsèques, vulnérabilités spécifiques et potentiels de contamination

Pour assurer la lisibilité de la méthode, ainsi qu'une flexibilité suffisante pour pouvoir décrire l'ensemble des situations possibles sur le territoire hexagonal, il nous a semblé nécessaire de dépasser le croisement global [Vulnérabilité intrinsèque * Pression phytosanitaire \Rightarrow Risque de contamination] du diagnostic régional Corpen. La décomposition en croisement de variables plus précis, constituant le « squelette » d'ARPEGES, est schématisée sur la Figure 4.

La méthode ARPEGES consiste d'abord en une évaluation de la **vulnérabilité dite « intrinsèque »** car liée au milieu physique (modifiée ou non par les aménagements anthropiques que sont par exemple les réseaux de drainage ou la végétalisation des berges). La vulnérabilité intrinsèque est déclinée selon les **trois grandes voies de transfert** hydrique contribuant à la contamination des eaux de surface : **ruissellement, écoulements de subsurface et drainage agricole**. Les écoulements de subsurface correspondent aux écoulements saturés qui peuvent advenir quand une couche de moindre perméabilité dans le sol limite la percolation de l'eau en profondeur, et conduit à des écoulements latéraux, suivant la pente de cette rupture de perméabilité. Ces écoulements sont en général lents et leur occurrence s'accompagne souvent de saturation en eau du profil de sol, qui peut conduire à la mise en place d'un drainage par tuyaux enterrés. L'occurrence et l'importance de ces différents types d'écoulement sont décrites, pour chaque situation, en fonction notamment des caractéristiques du sol et du sous-sol, de la présence ou non de protection des cours d'eau (ripisylve et zones tampons enherbées), ainsi que des caractéristiques de la zone climatique considérée qui modulent les transferts d'eau et de produits phytosanitaires (cumuls de pluie, fréquence et intensité des pluies). Ces vulnérabilités sont décrites pour deux saisons, nappe haute et nappe basse (cf 3.2.3.1).

Ensuite, on évalue la **vulnérabilité dite « spécifique »**, en combinant ces vulnérabilités intrinsèques avec les propriétés des substances actives : la vulnérabilité intrinsèque va plus ou moins s'exprimer en fonction des propriétés de la substance active considérée. Les propriétés prises en compte sont celles classiquement considérées dans les méthodes d'évaluation de risque : la demi-vie (DT50) et le coefficient d'adsorption sur le carbone organique (Koc). Dans des conditions de transfert similaires, les substances ne seront pas mobilisées de la même façon suivant leurs propriétés : au cours d'un événement pluvieux, une substance active s'adsorbant fortement sera significativement mobilisée si le milieu est vulnérable au ruissellement **et** érosif ; une substance s'adsorbant peu sera de toute façon fortement mobilisée par voie hydrique si le ruissellement est important, que le milieu soit érosif ou non. De même, dans un milieu vulnérable à une contamination via les écoulements de subsurface, les molécules rémanentes seront plus problématiques que les autres.

Enfin, le potentiel de contamination résulte du croisement entre la vulnérabilité spécifique et les quantités de substances utilisées (que l'on définit comme la « pression phytosanitaire ») : la probabilité de transfert au cours d'eau sera d'autant plus forte que la substance est utilisée, bien que la relation ne soit pas linéaire.

Comme le suggère le schéma de la Figure 4, des distinctions de saison (nappe haute/nappe basse) et de type de transfert (lents/rapides) sont réalisées pour tout ou partie des variables calculées par ARPEGES. Les paragraphes du 3.2.3 détaillent la signification et les particularités de calcul de chacune.

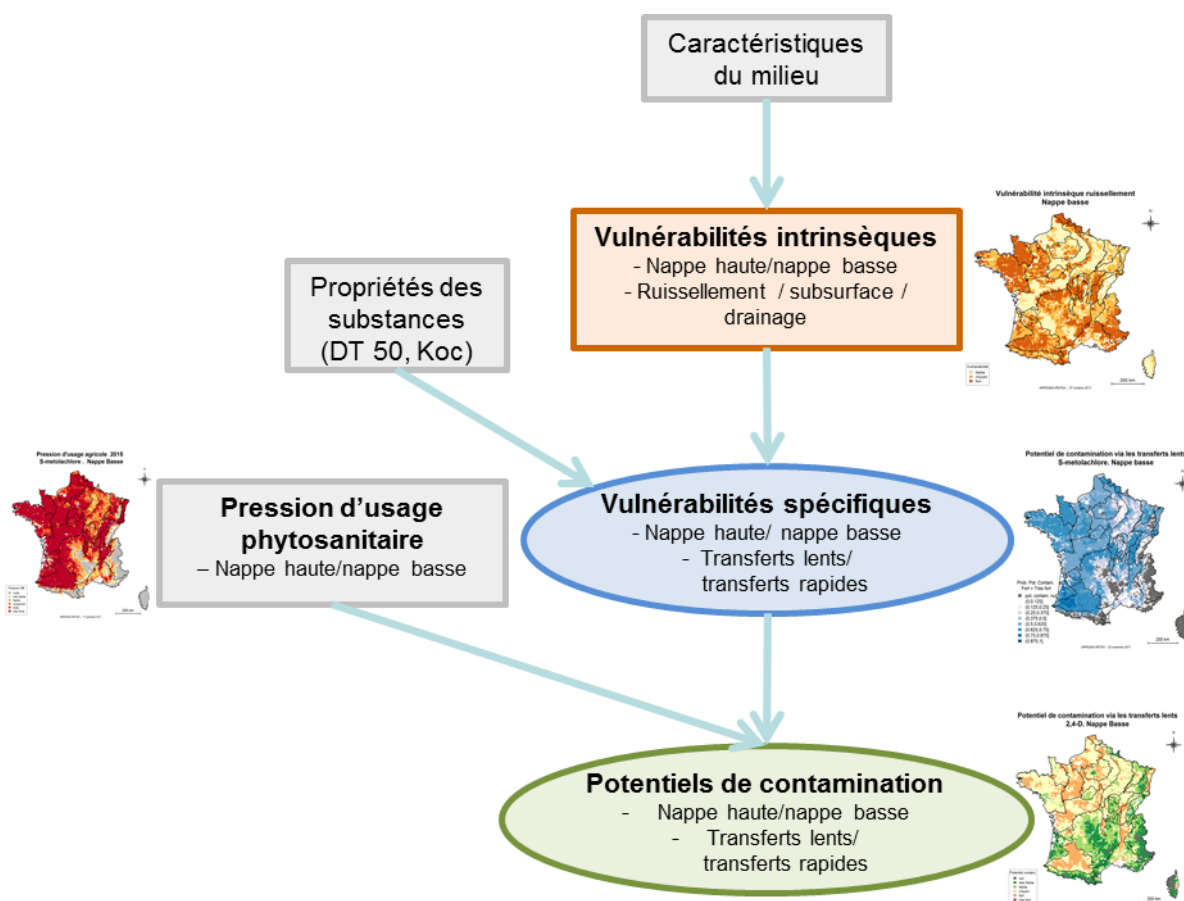


Figure 4 : Schéma très simplifié du modèle ARPEGES (les cadres gris correspondent aux données d'entrée du modèle, les cadres de couleur correspondent aux variables calculées par le modèle)

3.2.3. Résultat final : obtention de 4 types de potentiels de contamination

3.2.3.1. Notion de saison dans ARPEGES :

La méthode étant basée en partie sur une caractérisation des voies de transfert de l'eau, la vulnérabilité intrinsèque de chaque type d'écoulement, et par conséquent la vulnérabilité spécifique et le potentiel de contamination, sont déclinés en deux saisons, afin de tenir compte du fonctionnement différencié des bassins versants entre **période de nappe basse** (printemps-été-début d'automne) **et de nappe haute** (fin d'automne-hiver). Cette différenciation est en lien notamment avec la recharge des nappes, le niveau de remplissage de la réserve utile des sols et les caractéristiques des pluies.

Ces qualificatifs renvoient à des fonctionnements hydrologiques pour les écoulements vers les eaux de surface, et **ne font pas référence à la prise en compte de transferts d'eau vers le compartiment souterrain**. Les nappes hautes ou basses auxquelles il est fait référence sont des nappes temporaires superficielles qui se forment dans certaines conditions de sols et de saturation en période hivernale. En nappe haute, le drainage par drains enterrés ou dans les fossés est actif, les sols sont saturés et la réserve utile est pleine ou quasiment pleine. La distinction en deux saisons n'est pas apparente dans

les données d'entrée des vulnérabilités intrinsèques, mais intervient dans le cœur du modèle réalisé par les experts (cf3.3.2).

La pression phytosanitaire se décline également en ces deux mêmes saisons, afin de pouvoir mettre en face de chaque vulnérabilité spécifique une quantité de substance active appliquée et potentiellement transférée sur la même période. Cette dichotomie est à renseigner par l'utilisateur, mais elle peut poser des questions méthodologiques (applications en frontière de période) : des indications ont été structurées pour aider à estimer des coefficients de saisonnalité par saison (cf 3.4.1.2 et Annexe 1).

Une subtilité existe dans les variables croisées pour obtenir les potentiels de contamination nappe haute/nappe basse (Figure 5) :

- Le potentiel de contamination nappe basse est issu du croisement entre la **vulnérabilité spécifique nappe basse** et la **pression phytosanitaire nappe basse**.
- Le potentiel de contamination nappe haute est issu du croisement entre la **vulnérabilité spécifique nappe haute**, la **pression phytosanitaire nappe haute et nappe basse**. La pression phytosanitaire nappe basse vient dans ce cas seulement augmenter légèrement le potentiel de contamination nappe haute, afin de rendre compte du possible entraînement en nappe haute de résidus de substance appliquée en nappe basse.

Les limites de ces périodes peuvent en réalité varier en fonction de l'année (plus ou moins pluvieuse) et de la zone climatique considérée (par exemple, l'automne est plus tardif et le printemps plus précoce dans la zone océanique que dans la zone continentale Est). Cependant, les dates bornant ces deux périodes ne sont pas formalisées dans le modèle. Comme le raisonnement est basé sur le fonctionnement hydrologique des sols, et que la période d'application des produits phytosanitaires dépend également en général de l'état hydrique du sol, cette façon de procéder ne pose pas de problème. Il s'agit de parler, en ce qui concerne la pression d'usage phytosanitaire, d'applications en période de saturation des sols ou d'applications en période de non-saturation.

3.2.3.2. Distinction des vulnérabilités spécifiques en deux types de contamination

Dans la même logique, on distingue la vulnérabilité spécifique (et par conséquent le potentiel de **contamination**) **par transferts rapides** de la vulnérabilité spécifique (et potentiel de contamination) **par transferts lents**, afin de mieux rendre compte du fonctionnement hydrologique du milieu. En effet, les processus en jeu ne sont pas les mêmes. Les transferts rapides adviennent **pendant ou juste après un événement pluvieux**, et sont générés par ruissellement et drainage en période d'écoulement intense. Les transferts lents, quant à eux, adviennent **en période de récession**⁵, et englobent les 3 types d'écoulements : ruissellement, drainage et écoulements de subsurface.

Au-delà de la prise en compte de tout ou partie des 3 types d'écoulements, les croisements entre variables varient pour ces différents types de potentiel, pour rendre compte au mieux des processus et interactions à l'œuvre selon les cas (Figure 5):

⁵ Période suivant un épisode pluvieux, et correspondant au ressuyage progressif des sols.

- la vulnérabilité spécifique aux transferts lents est issue du croisement entre **le cumul des pluies**, les propriétés de la substance active, la vulnérabilité intrinsèque au ruissellement, au drainage et aux écoulements de subsurface.
- la vulnérabilité spécifique aux transferts rapides est issue du croisement entre **la fréquence des pluies, la matière organique des sols**, les propriétés de la substance active, la vulnérabilité intrinsèque au ruissellement et au drainage.

Les potentiels de contamination par transferts rapides et par transferts lents sont obtenus respectivement à partir de la vulnérabilité spécifique aux transferts rapides et aux transferts lents.

3.2.3.3. Synthèse des différentes variables calculées par ARPEGES

Le modèle semi-détaillé d'ARPEGES, développant les différentes distinctions présentées dans les deux paragraphes précédents, est schématisé sur la Figure 5 pour chaque saison (a - nappe haute et b - nappe basse). Pour des raisons de simplicité de lecture, les catégories « Caractéristiques du milieu », « Zonage climatique », « Cumul de pluies » et « Matière organique du sol » apparaissant sur la Figure 5 étaient englobées dans la catégorie « Caractéristiques du milieu » de la Figure 4.

Pour chaque bassin versant local de masse d'eau, ARPEGES calcule donc :

- 6 vulnérabilités intrinsèques (3 types d'écoulement et 2 saisons), dont les résultats spatialisés par principe ne dépendent pas de la substance active étudiée (cf. 3.5.3);
- 4 vulnérabilités spécifiques (transferts lents et rapides et 2 saisons) : les résultats spatialisés de chaque vulnérabilité spécifique sont identiques pour chaque groupe de substances types (combinaison Koc/DT50, cf. 3.5.3)
- 4 potentiels de contamination (transferts lents et rapides et 2 saisons)

NB : même si les 4 combinaisons sont calculées par le modèle, on notera que les transferts lents en période de nappe basse sont moins probables que les trois autres combinaisons.

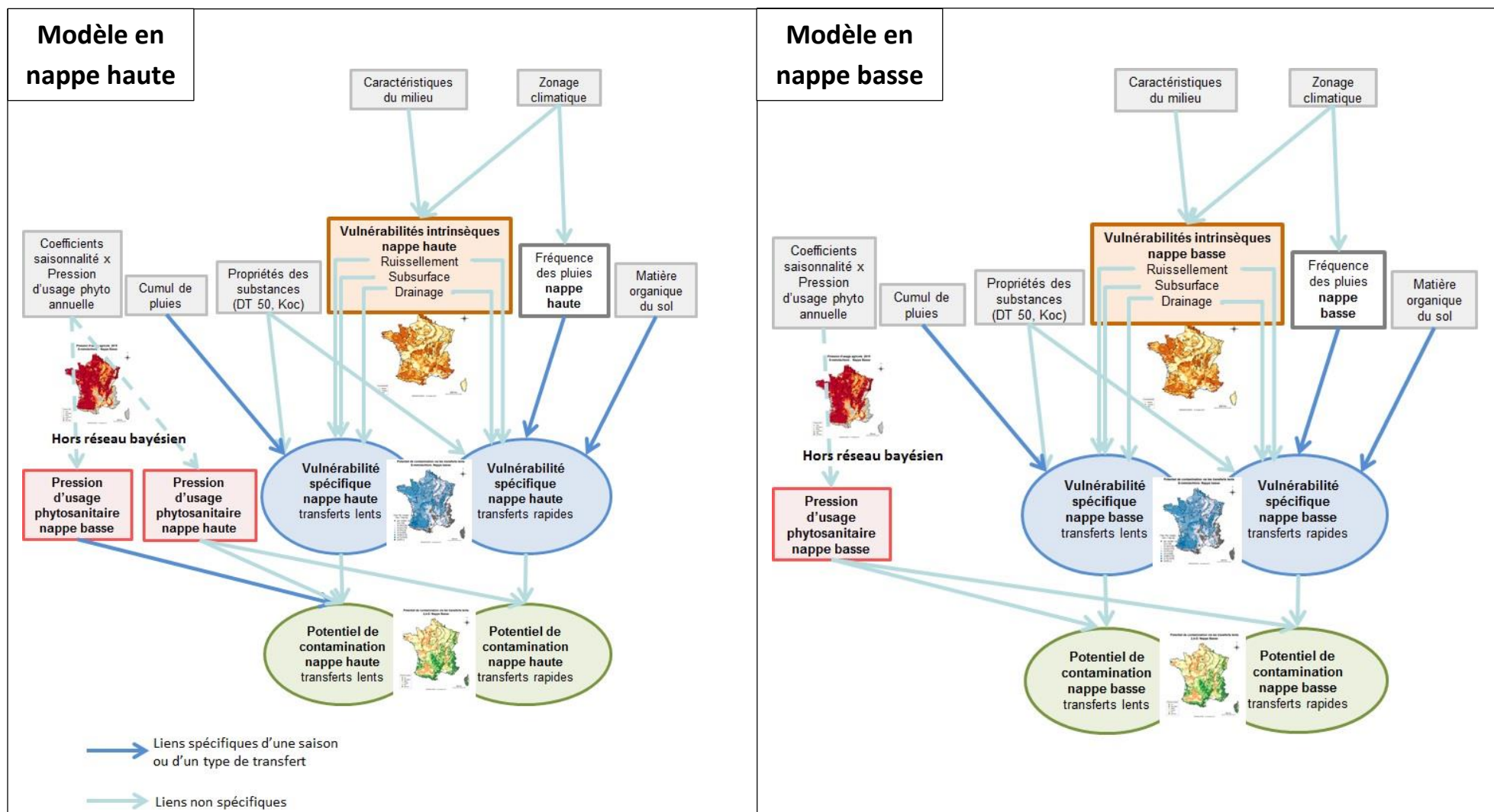


Figure 5 : Schéma simplifié du modèle ARPEGES (à droite : 5a - nappe haute ; à gauche : 5b - nappe basse)

3.2.4. Approche probabiliste du modèle

A cette échelle, le croisement des données de façon purement mécaniste n'a pas été retenu, compte tenu de la complexité des processus en jeu et de leurs interactions, mais aussi de la nature des données disponibles à cette échelle. La complexité des croisements a donc pu être rendue en ayant recours à la méthode des réseaux bayésiens. Cette approche probabiliste permet de tenir compte de l'incertitude sur les données et de formuler des connaissances sur les processus dominants conditionnellement à certaines variables (cf. partie 4.3).

Par exemple : « si les sols sont très battants sur ce bassin versant et que les bords de rivière ne sont pas protégés par une ripisylve ou des zones tampons enherbées, alors il y a de fortes chances que le ruissellement hortonien advienne fréquemment et atteigne la rivière ».

Les réseaux de probabilité bayésiens reposent sur une représentation graphique (cf. 3.3.1) de variables aléatoires (les nœuds) liées entre elles par des relations de cause à effet (les flèches) probabilisées (cf. 3.3.2).

Ce sont des variables aléatoires et leurs valeurs ne sont donc pas considérées comme une quantité unique certaine ; on décrit au contraire leur distribution de probabilités, c'est-à-dire la probabilité qu'a cette variable de prendre chacune de ses valeurs potentielles (ses états), ou d'appartenir à un intervalle de valeur donné, dans le cas des variables continues.

La création d'un réseau bayésien nécessite plusieurs étapes :

- **Choix des variables d'intérêt (cf. 3.3.1.1).** *Par exemple, la réserve utile, la pression phytosanitaire, la vulnérabilité spécifique aux transferts lents ou encore le potentiel de contamination par les transferts lents.*
- **Choix des états qu'elles peuvent prendre (cf. 3.3.1.2).** Il peut s'agir de valeurs discrètes qualitatives ou quantitatives, ou bien d'intervalles de valeurs pour les variables continues (définition des limites de classe dans ce cas). *Par exemple :*
 - *la vulnérabilité spécifique aux transferts lents peut être faible, moyenne ou forte ;*
 - *la pression phytosanitaire peut être très faible, faible, moyenne, forte ou très forte. Les limites de classe correspondantes qui ont été choisies sont [0-0,1 g/ha de bassin versant] ; [0,1-1 g/ha] ; [1-5 g/ha] ; [5-10 g/ha] et [> 10 g/ha].*
 - *le potentiel de contamination par transferts lents peut être faible, moyen ou fort.*
- **Etablissement des relations entre les variables et expression des tables de probabilités conditionnelles pour chaque nœud (cf. 3.3.1.3 et 3.3.2).** *Par exemple :*
 - *Relation entre variables : Le risque de contamination par transferts lents dépend de la vulnérabilité spécifique aux transferts lents et de la pression d'usage phytosanitaire.*
 - *Probabilités conditionnelles pour cette relation : « Si la vulnérabilité et la pression sont faibles, alors le risque de contamination aura une probabilité de 100 % d'être faible. » ; « Si la pression est faible et la vulnérabilité est moyenne, alors le risque de contamination aura une probabilité de 70 % d'être faible et de 30 % d'être moyen. » ; et ainsi de suite pour chaque combinaison possible.*

Compte tenu de la méthode de construction de ces tables, qui intègre à la fois les relations de causalités connues et l'incertitude que l'on peut avoir sur les données d'entrée (considérées comme des variables aléatoires), un réseau bayésien permet d'assortir les résultats d'un niveau de confiance, permettant de qualifier l'incertitude sur le résultat acquis (cf. 3.5.1.3).

La structure de modèle décrite aux §3.2.2 et 3.2.3 a donc été traduite en un réseau de probabilités bayésien. Dans le cadre de la méthode ARPEGES, la sélection des variables pertinentes, la discrétisation des variables continues, l'établissement des relations entre les variables et des tables de probabilité conditionnelle associées ont été effectués par expertise. Le processus s'est avéré itératif :

- pour finaliser la structure du modèle, finalement assez éloigné du schéma initial Corpen, car les réseaux bayésiens permettent des nuances que ne permettait pas la stricte approche déterministe utilisée dans la démarche Corpen,
- pour fixer les limites des classes d'état et les valeurs des tables de probabilité, en vérifiant que les cartes obtenues pour les variables intermédiaires (par exemple vulnérabilité intrinsèque à un type de transfert pour une saison donnée) correspondaient bien à la connaissance que les experts avaient du milieu.

Ce travail a été mené en tenant compte des données disponibles pour pouvoir appliquer la méthode dès à présent.

L'écriture du modèle sous forme de graphe, l'entrée des tables de probabilités conditionnelles ainsi que le calcul des probabilités résultantes pour les variables de sortie, ont d'abord été réalisés à l'aide du logiciel dédié à l'inférence dans les réseaux de probabilité bayésiens NETICA. En parallèle de cette étude, un travail de développement d'une interface sous logiciel libre a été réalisé (Mellac et Piffady, 2018) et rend ainsi le modèle accessible à tous.

3.3. Construction du réseau bayésien d'ARPEGES

Cette partie permet de connaître précisément la structure du modèle ARPEGES, préalable nécessaire pour bien comprendre le mode de calcul des sorties, qui est traité dans la partie 3.5. Cette partie 3.3 n'aborde pas le détail des données d'entrée, pour lesquelles une partie à part entière est consacrée (partie 3.4).

Cette partie est structurée selon les 4 étapes de la construction d'un réseau bayésien : les 3 premières concernent la structure graphique, la 4^{ème} concerne la structure probabiliste.

3.3.1. Construction de la structure graphique du réseau

Un encadré « Les fondements mathématiques 1 » est dédié au vocabulaire spécifique et aux concepts mathématiques de la structure graphique des réseaux bayésiens.

3.3.1.1. 1^{ère} étape : Choix des variables d'intérêt

Le choix des variables pertinentes à intégrer au réseau a été fait par expertise. Ce choix n'a pu se faire qu'en tenant compte des sources de données disponibles, pour que la méthode soit applicable dès maintenant.

Le tassement et la couverture des sols, bien qu'influençant le ruissellement hortonien (ruissellement se produisant lorsque l'intensité de la pluie est supérieure au taux d'infiltration du sol), n'ont pas été retenus pour un travail à cette échelle, d'une part à cause de la complexité de l'estimation de ces paramètres (modélisation ou expertise à partir de l'occupation du sol, du type de sol et du climat), et d'autre part parce qu'ils varient beaucoup d'une année à l'autre selon les conditions climatiques.

C'est également le cas de la pente des versants. Bien que ce soit un processus important qui influe sur le ruissellement à l'échelle locale, cette variable n'a pas été retenue car trop complexe à intégrer dans un tel modèle s'appliquant à l'échelle du bassin versant (quelle valeur considérer ? quelle donnée mobiliser ? la pente moyenne a-t-elle un sens à l'échelle d'un bassin versant ? prise en compte de la concavité des versants ?).

Les fondements mathématiques 1 : Structure graphique des réseaux bayésiens

L'écriture d'un réseau de probabilités bayésien repose sur une représentation graphique schématique du système et de son fonctionnement. Ces modèles s'appuient directement sur la théorie des graphes. On distingue deux principaux éléments graphiques

► Les nœuds :

Les variables d'intérêt dans le fonctionnement du système sont représentées sous la forme de boîtes, appelées « nœuds ». Ce sont des variables aléatoires et leur valeur n'est donc pas considérée comme une quantité fixée, mais on décrit au contraire leur distribution de probabilités, c'est-à-dire la probabilité qu'a cette variable de prendre chacune de ses valeurs potentielles, ou d'appartenir à un intervalle de valeur donné dans le cas des variables continues.

Les réseaux de probabilités bayésiens ne peuvent utiliser directement des variables continues, et celles-ci doivent être discrétisées. Cette étape consiste en la description d'intervalles de valeurs, chacun correspondant au final à une classe. On peut construire ces intervalles de façon naïve, en découpant l'espace en intervalles de même taille, par exemple ; ou bien construire ces intervalles en faisant des hypothèses sur le fonctionnement du système, on décrit alors des intervalles de valeurs pour lesquelles le comportement du système est considéré comme stable.

► Les flèches :

Pour décrire des relations causales entre les nœuds, on utilise un système de flèches. Un nœud duquel part une flèche est appelé un nœud parent, un nœud sur lequel pointe une flèche est appelé un nœud enfant. Tracer une flèche d'un nœud parent vers un nœud enfant sous-entend que l'on considère qu'il existe un lien de causalité du nœud parent vers le nœud enfant, autrement dit, on considère que les valeurs possiblement prises par le nœud enfant dépendent des valeurs prises par le nœud parent.

Les nœuds parents et enfants retenus dans la méthode ARPEGES sont représentés sur la Figure 6.

⇒ **Les nœuds « parents »** (données d'entrée) peuvent être regroupés en sous-groupes thématiques :

- Données pédologiques (fait partie de la catégorie « Caractéristiques du milieu » sur la Figure 5) :
 - l'IDPR (Indice de Développement et de Persistance des Réseaux (Van Genuchten, Leij et al. 1991)), qui donne une indication sur le partage de l'eau entre réseau de surface et eaux souterraines et donc sur l'importance des écoulements de surface et proche subsurface ;
 - la sensibilité des sols à la formation d'une croûte de battance, qui augmente le potentiel de ruissellement hortonien ;

- la présence d'hydromorphie et de la valeur de la réserve utile, qui influent sur le ruissellement par saturation ;
- le taux de matière organique dans les sols, qui interagit avec les substances s'adsorbant fortement.
- Données sur les aménagements des sols et du paysage (fait partie de la catégorie « Caractéristiques du milieu » sur la Figure 5) :
 - La part de la surface agricole utile du bassin versant drainée par un réseau de tuyaux enterrés;
 - la présence de ripisylves ou de bandes enherbées, qui peuvent atténuer le ruissellement avant son arrivée au réseau hydrographique.
- Données climatiques :
 - la zone climatique, qui influence la fréquence et l'intensité des pluies ;
 - le cumul des pluies ;
 - la saison (nappe haute ou basse), qui influence tous les nœuds enfants du réseau bayésien ainsi que la fréquence des pluies .
- Données relatives à la substance active étudiée :
 - le Koc, qui renseigne sur la capacité de la substance à s'adsorber sur les particules de sol, et la DT50, sur sa rémanence.
 - la quantité de substance active appliquée en nappe haute et en nappe basse par unité de surface de BVME (appelée aussi pression d'usage phytosanitaire nappe haute et nappe basse).

On insiste ici sur l'importance de rapporter la quantité de substance active à la surface du BVME : en effet, cela permet de prendre en compte l'effet de dilution des bassins versants. Par exemple, entre deux bassins versants de tailles différentes contenant la même surface traitée par une substance active donnée, c'est le bassin versant de plus petite taille qui présentera la plus forte pression d'usage de cette substance, car la quantité appliquée sera moins diluée.

⇒ **Les nœuds « enfants »** ont déjà été détaillés dans la partie 3.2 précédente :

- Vulnérabilités intrinsèques pour les trois types d'écoulements (ruissellement, drainage agricole, écoulement de subsurface) ;
- Vulnérabilités spécifiques aux transferts lents et rapides ;
- Potentiels de contamination par transferts lents et rapides.

La distinction en deux saisons étant déjà intégrée dans le nœud parent « saison », il n'y a pas de nœud enfant décliné pour chaque saison dans la structure graphique du réseau bayésien d'ARPEGES. La distinction en saison se fait directement dans les tables de probabilités (cf 3.3.2).

Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 » 1 : Choix des variables d'intérêt

- ARPEGES ne calcule plus un potentiel de contamination pour un « cocktail » de substances actives, mais pour chaque substance active. Cela a donc changé les données d'entrée pour les variables « DT50 » et « Koc » : ce ne sont plus des valeurs moyennes pondérées selon les usages de chaque substance à l'échelle des HER, mais la valeur de référence de chaque variable pour la substance active étudiée. Ce changement présente l'intérêt de remonter plus facilement aux causes du potentiel de contamination, et permet de tester la cohérence des sorties d'ARPEGES à des données de surveillance des cours d'eau.
- Plusieurs variables ont été supprimées du modèle, pour différentes raisons relatives aux liens entre variables :
 - la solubilité de la substance active (nœud parent) : à l'exception du glyphosate, cette caractéristique joue dans le même sens que le Koc, et les connaissances actuelles ne sont pas suffisantes pour intégrer la conjonction de ces deux variables dans le modèle de façon pertinente.
 - l'aléa érosif (nœud parent) : le transfert des substances actives s'adsorbant fortement se fait surtout sur des particules fines ou colloïdales, indépendamment de l'érosion des éléments grossiers.
 - la densité de réseau hydrographique (nœud parent) et la vulnérabilité intrinsèque à la dérive (nœud enfant) : la limitation du risque lié à la dérive dépend de l'application de la réglementation (utilisation de buses anti-dérive, zones non-traitées, dispositifs végétalisés permanents) dont on ne dispose pas d'une connaissance suffisante à l'échelle nationale.

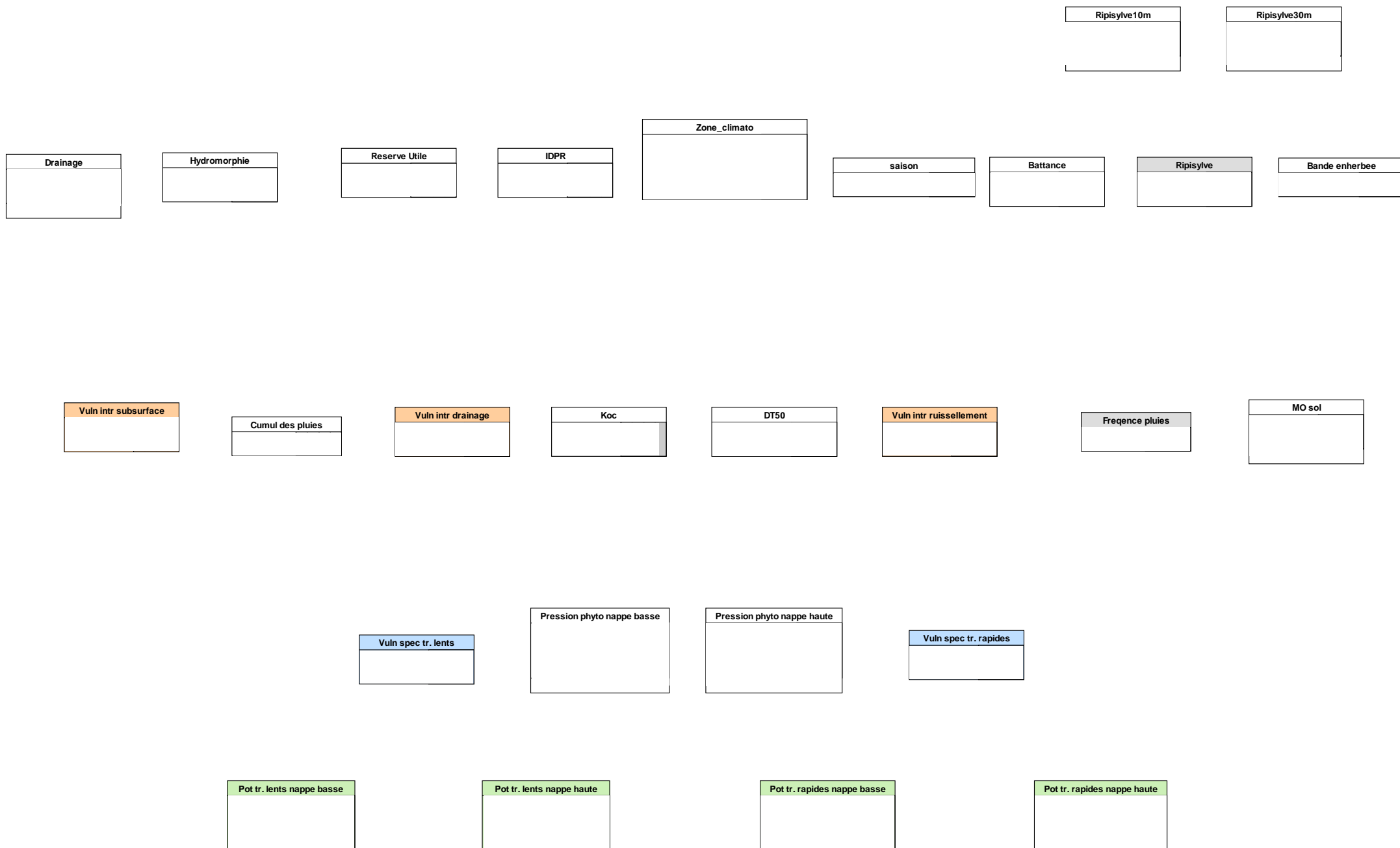


Figure 6 : Construction de la structure graphique d'ARPEGES - Choix des nœuds parents et enfants

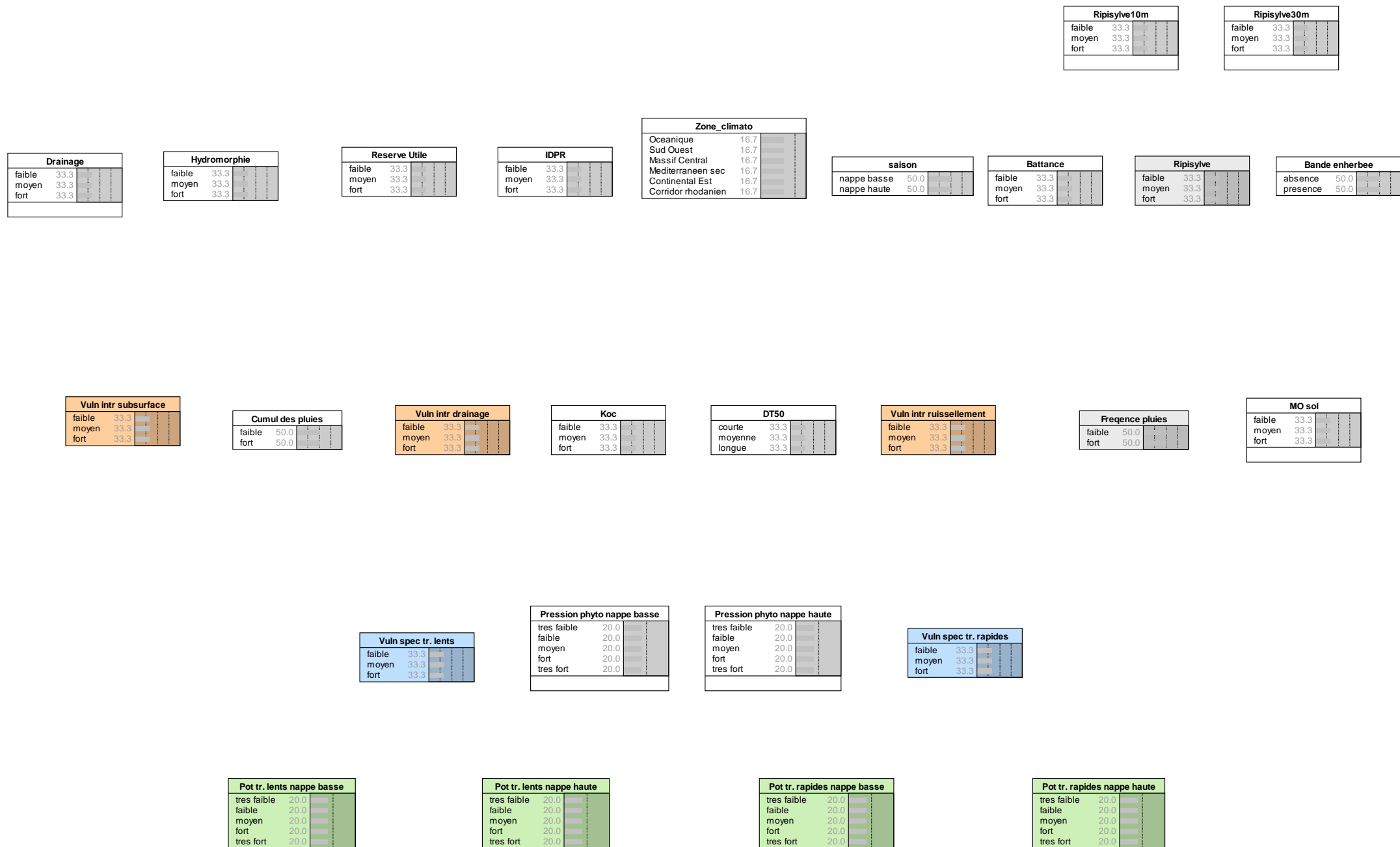


Figure 7 : Construction de la structure graphique d'ARPEGES - Choix des états pris par les nœuds parents et enfants

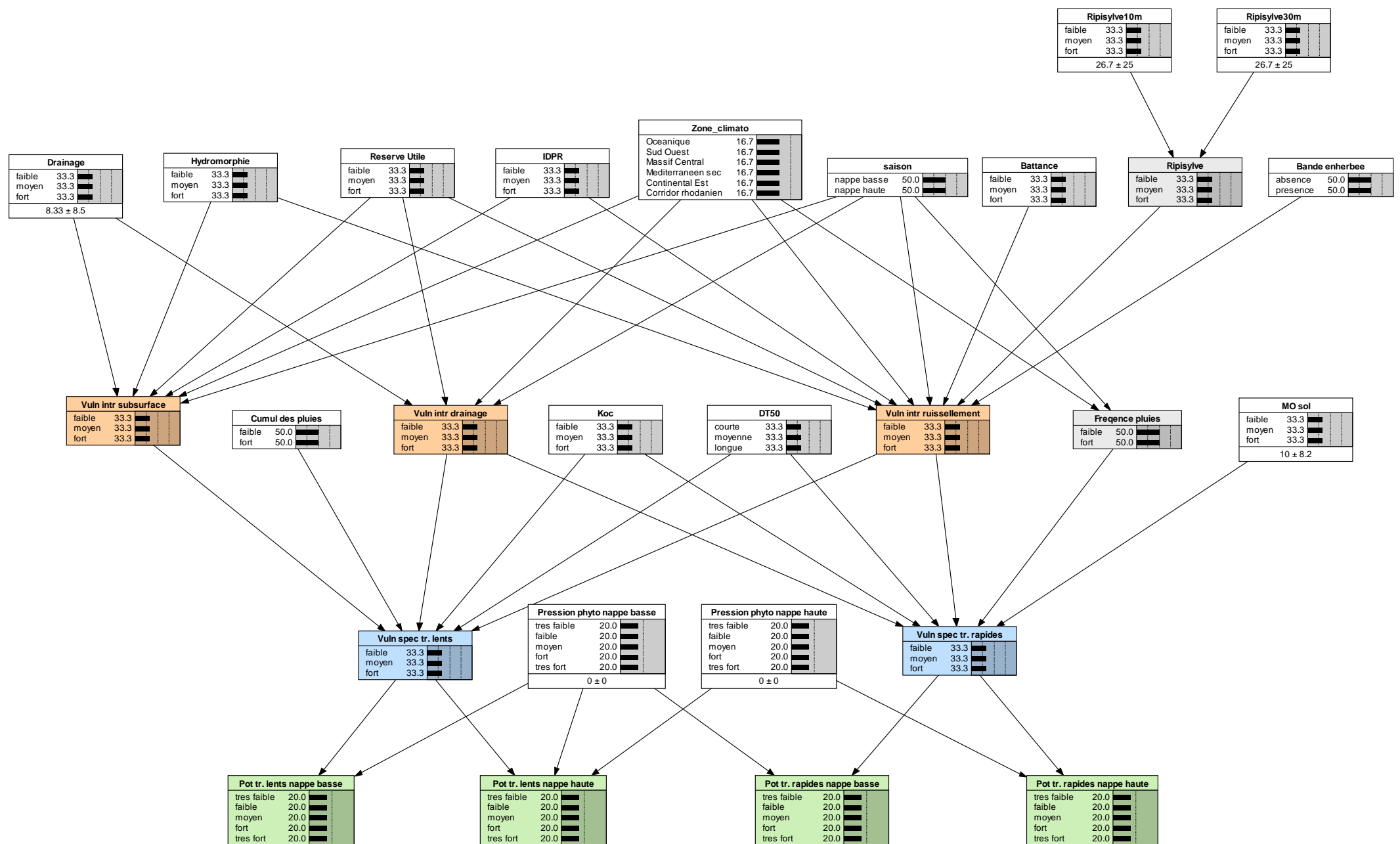


Figure 8 : Construction de la structure graphique d'ARPEGES - Etablissement des liens entre les nœuds parents et enfants

L'apparence du réseau bayésien en étape 2 après choix des états des nœuds parents et enfants est présentée dans la Figure 7. On voit qu'à ce stade, tous les états de toutes les variables sont équiprobables.

⇒ **Nœuds « parents » :**

Les données en entrée peuvent être de nature discrète ou continue. Des intervalles de valeurs sont donc définis après choix des états pris par la variable si celle-ci est continue. Les états et intervalles de valeurs sont très liés à la nature des données fournies en entrée, et sont éventuellement à revoir dans le cas d'une actualisation de données d'entrée. Pour plus de précisions sur les données d'entrées, voir la partie 3.43.4.

⇒ **Nœuds « enfants » :**

- Les vulnérabilités intrinsèques et spécifiques ont trois états possibles : faible, moyen et fort.
- Les potentiels de contamination ont cinq états possibles : très faible, faible, moyen, fort et très fort.

Le réseau bayésien tel qu'obtenu en étape 3 (après établissement des liens entre variables) est représenté sur la Figure 8. On remarque qu'un nœud parent peut être relié à plusieurs nœuds enfants : cela vaut pour les variables d'entrée (nœuds parents uniquement), mais aussi pour les variables « intermédiaires » (vulnérabilités intrinsèques notamment), qui sont à la fois des nœuds enfants et parents. On note aussi que les flèches n'affectent pas à ce stade l'équiprobabilité des états des variables : c'est une fois le renseignement des tables de probabilités effectué que l'on peut voir une différence dans les probabilités des états des variables (voir §3.3.2 et Figure 15). Les rubriques de ce paragraphe expliquent les liens entre chaque nœud enfant et les nœuds parents associés. Pour plus d'informations sur chaque variable d'entrée, voir partie 3.4.

Vulnérabilité intrinsèque au ruissellement

Le ruissellement de surface peut être séparé en deux types, le ruissellement par saturation et le ruissellement hortonien.

Le ruissellement par saturation a lieu quand la surface du sol est saturée en eau (**hydromorphie**). Une faible **réserve utile** en eau des sols et un **cumul de pluies** important sont des facteurs de risque.

Le ruissellement hortonien est dû à un défaut d'infiltration de l'eau dans le sol : l'**intensité de la pluie** est supérieure à la capacité d'infiltration du sol, à cause de la formation d'une croûte de **battance**, d'un tassement du sol ou si la pluie est très intense (pluies cévenoles par exemple).

Les paramètres déjà cités influencent la genèse du ruissellement, mais celui-ci peut être atténué avant son arrivée au réseau hydrographique, soit grâce à la présence d'une zone tampon, soit grâce à une meilleure capacité d'infiltration des sols. Dans notre méthode, nous n'intégrons comme zones tampons que les ripisylves et les bandes enherbées, bien que d'autres types existent (zones humides par exemple). Nous avons posé l'hypothèse que les bandes enherbées et les ripisylves ne sont pas court-circuitées par le ruissellement car c'est une information impossible à connaître à cette échelle.

Thématique de données d'entrée	Nœud parent	Intitulé de la donnée disponible	Type de donnée	Etats	Seuils de discrétisation	Probabilité de l'état
Données pédologiques	IDPR	idem	continue	faible	<800	100 % pour l'état résultant de la discrétisation
				moyen	800 < x < 1200	
				fort	>1200	
	Battance des sols	Sensibilité des sols à la formation d'une croûte de battance	discrète	faible		% de la surface du BV en faible
				moyen		% de la surface du BV en moyen
				fort		% de la surface du BV en fort
	Hydromorphie des sols	Présence de sols hydromorphes	discrète	faible		100 % pour l'état résultant de la discrétisation
				moyen		
				fort		
	Réserve utile	idem	discrète	faible		% de la surface du BV en faible
				moyen		% de la surface du BV en moyen
				fort		% de la surface du BV en fort
	Matière organique du sol	Teneur du sol en carbone organique	continue	faible	< 10 g/kg	100 % pour l'état résultant de la discrétisation
				moyen	10 < x < 20 g/kg	
				fort	> 20 g/kg	
Données sur les aménagements des sols et du paysage	Drainage agricole	Pourcentage de la SAU drainée	continue	faible	< 5 % de la SAU	100 % pour l'état résultant de la discrétisation
				moyen	5 < x < 20 %	
				fort	> 20 % de la SAU	
	Ripisylve 10 m	Surface en forêt dans la bande de 10 m au bord du cours d'eau	continue	faible	< 20 % de la surface de la bande	100 % pour l'état résultant de la discrétisation
				moyen	20 < x < 60 %	
				fort	> 60 % de la surface de la bande	
	Ripisylve 30 m	Surface en forêt dans la bande de 30 m au bord du cours d'eau	continue	faible	< 20 % de la surface de la bande	100 % pour l'état résultant de la discrétisation
				moyen	20 < x < 60 %	
				fort	> 60 % de la surface de la bande	
	Bande enherbée	Pourcentage de linéaire de cours d'eau devant être protégé	discrète	présence		% de cours d'eau avec un nom
				absence		% de cours d'eau sans nom (= 100-présence)
Données climatiques	Zone climatique	idem	discrète	Océanique		100 % pour la zone climatique du BV
				Sud-Ouest		
				Massif central		
				Méditerranéen sec		
				Continental Est		
				Corridor rhodanien		
	Cumul de pluie	idem	discrète	faible		100 – (Cumul mensuel * 100 / 300)
				fort		Cumul mensuel * 100 / 300
	Saison	Pas de donnée à proprement parler	discrète	nappe basse		100 % pour la saison étudiée
				nappe haute		
Données sur la substance active	Koc	idem	continue	faible	< 100 L/kg	100 % pour l'état résultant de la discrétisation
				moyen	100 < x < 1000	

étudiée				fort	> 1000 L/kg	100 % pour l'état résultant de la discrétisation
	DT50	idem	continue	courte	< 8 jours	
				moyenne	8 < x < 30	
				longue	> 30 jours	
	Pression d'usage phytosanitaire nappe haute	Quantité de substance active appliquée par ha de BVME en nappe haute	continue	très faible	< 0,1 g/ha de BVME	100 % pour l'état résultant de la discrétisation
				faible	0,1 < x < 1	
				moyen	1 < x < 5	
				fort	5 < x < 10	
				très fort	> 10 g/ha de BVME	
	Pression d'usage phytosanitaire nappe basse	Quantité de substance active appliquée par ha de BVME en nappe basse	continue	très faible	< 0,1 g/ha de BVME	100 % pour l'état résultant de la discrétisation
				faible	0,1 < x < 1	
				moyen	1 < x < 5	
				fort	5 < x < 10	
				très fort	> 10 g/ha de BVME	

Tableau 3 : Etats des nœuds parents et seuils de discrétisation

Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 » 2 : Choix des états et seuils de discrétisation

- Le nombre d'états et la valeur des seuils des pressions d'usage phytosanitaire ont été modifiés :

Dans la méthode ARPEGES de 2013, la donnée utilisée pour la pression phytosanitaire était le RICA car on étudiait un cocktail de substances actives. On appliquait alors des seuils fixes, définis par expertise (<10€/ha, de 10 à 30€/ha, >30€/ha).

Dans une étude intermédiaire datant de 2016 et commanditée par l'Anses, l'adaptation d'ARPEGES pour une évaluation substance par substance avait déjà été effectuée. Cependant, il n'y avait encore que trois classes pour les pressions d'usage phytosanitaire, et les seuils des classes faible/moyen/fort des pressions d'usage phytosanitaire étaient calculées sur la base des 33 et 66^{ème} centiles des valeurs des quantités appliquées par saison de la substance active étudiée (en g/ha de bassin versant) sur l'ensemble des BVME français, pour les deux saisons nappe basse et nappe haute. On avait donc une équité-répartition des pressions d'usage avec 33% du nombre de BVME en classe faible, 33% en moyen et 33% en fort pour une substance donnée, et ceci pour toutes les substances actives étudiées, quel que soit le tonnage appliqué pour chaque molécule. Cette méthode avait pour intérêt de bien discriminer la pression d'usage entre les BVME.

Il était alors possible d'avoir des substances actives appliquées en très faibles quantités, et qui se retrouvaient dans ARPEGES sur un certain nombre de bassins en potentiel de contamination « fort », pour des quantités très inférieures aux quantités appliquées de substances actives très utilisées sur des bassins ressortant en potentiel de contamination « faible ». De plus, au-delà d'une impossibilité de comparer les résultats entre substances actives, cette façon de procéder impliquait une distorsion des principes du modèle, en prenant des seuils de pression « relatifs » à chaque substance active, alors que les tables de croisement pression d'usage/vulnérabilité spécifique du modèle ont été construites en raisonnant par rapport à une pression d'usage phytosanitaire « absolue ».

Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 » 2 : Choix des états et seuils de discrétisation (suite)

Cependant, la pression était considérée en termes relatifs : les seuils de pression d'usage phytosanitaire étaient propres à chaque substance, et étaient ainsi très différents selon les substances, comme on peut le lire dans le tableau suivant :

Substance active	Centile 33 des quantités appliquées par saison (g/ha de BV)	Centile 66 des quantités appliquées par saison (g/ha de BV)
2,4-D	0,5	3,1
2,4-mcpa	2,2	8,9
Aminotriazole	0,1	0,8
Bentazone	0,9	3,7
Boscalid	1,0	4,6
Chlortoluron	3,2	17,8
Diflufenicanil	0,3	2,1
Glyphosate	20,5	65,5
Metazachlore	0,5	3,1
Nicosulfuron	0,2	1,1
Pendimethaline	2,3	9,9
S-metolachlore	7,6	37,9

(doublon)

Pour pallier ces limites, et après consultation des parties prenantes du projet (Agences de l'eau, Ministère de l'environnement, AFB), le nombre d'états pour les pressions d'usage phytosanitaire est passé de trois à cinq, et des valeurs « absolues » de seuils (identiques pour toutes les substances actives) ont été établies en considérant l'amplitude des valeurs prises par l'échantillon des quinze substances actives étudiées dans le projet. CE choix présente entre autres l'avantage de permettre d'évaluer l'effet de l'évolution des substances entre deux états des lieux.

- Le nombre d'états des potentiels de contamination :

Profitant de la remise à plat des états de la pression d'usage phytosanitaire, et par conséquent des tables de probabilité des potentiels de contamination, le nombre de classes des potentiels de contamination est passé de trois à cinq également. Ceci présente l'intérêt de pouvoir faire correspondre les classes de potentiel de contamination avec des intervalles de concentrations observées moins grands. Le raisonnement de la mise à plat a reposé sur un maintien de la signification de la modalité « moyen » pour les anciennes et nouvelles tables, afin de pouvoir comparer les résultats obtenus entre les deux « versions » du modèle ARPEGES. Cependant, cet ajout de classes n'apporte pas plus de précision aux résultats car il n'est pas lié à un ajout de connaissances de base : il faut donc interpréter cette précision apparente avec précaution.

Les possibilités de l'interface R : Choix des états et seuils de discrétisation

Le nombre et l'intitulé des états des variables ne sont pas modifiables dans l'interface R, mais celle-ci permet de modifier les seuils de discrétisation pour les variables continues. Cette fonctionnalité est toutefois à utiliser avec précaution car les tables de probabilités ont été construites en référence aux seuils de discrétisation initiaux listés dans le Tableau 3 (cf §3.3.2).

Les paramètres choisis pour estimer la vulnérabilité intrinsèque au ruissellement sont donc :

- ▶ La présence de sols hydromorphes
- ▶ La réserve utile
- ▶ La sensibilité des sols à la formation d'une croûte de battance
- ▶ La saison (tout le modèle est décliné selon les 2 saisons définies)
- ▶ La zone climatique. Ce paramètre permet de prendre en compte de façon experte les paramètres climatiques qui entrent en jeu dans les phénomènes de ruissellement, notamment l'intensité, la fréquence et le cumul des pluies.
- ▶ L'IDPR ou Indice de Développement et de Persistance des Réseaux. Il donne une indication si la destination de l'eau est bien le réseau de surface ou si l'eau finit par s'infiltrer vers une ressource profonde.
- ▶ La présence d'une ripisylve
- ▶ La présence d'une bande enherbée

Le croisement de toutes ces variables peut être représenté avec les cartes des nœuds parents et enfants, comme sur la

Figure 9.

Vulnérabilité intrinsèque aux flux de subsurface

Bien que ces flux soient encore très peu quantifiés in situ, il nous a semblé important de les intégrer dans notre méthode en considérant que leur prise en compte pourrait être améliorée ultérieurement avec l'avancée des connaissances sur ces transferts. Les flux de subsurface sont favorisés par la présence d'un horizon ou d'un socle imperméable et/ou d'une faible **réserve utile**. La trace d'un engorgement dans le sol (**hydromorphie**) peut traduire la présence d'un horizon imperméable. Le **drainage** limite les flux de subsurface en désengorgeant les sols.

Les paramètres choisis pour estimer la vulnérabilité intrinsèque aux flux de subsurface sont donc (voir Figure 10):

- ▶ La présence de sols hydromorphes
- ▶ La réserve utile
- ▶ Le pourcentage de la SAU drainée
- ▶ L'IDPR (indice de développement et de persistance des réseaux). Il permet de confirmer la destination de l'eau : eaux souterraines (donc peu de flux de subsurface) et/ou réseau de surface.
- ▶ La saison

► La zone climatique

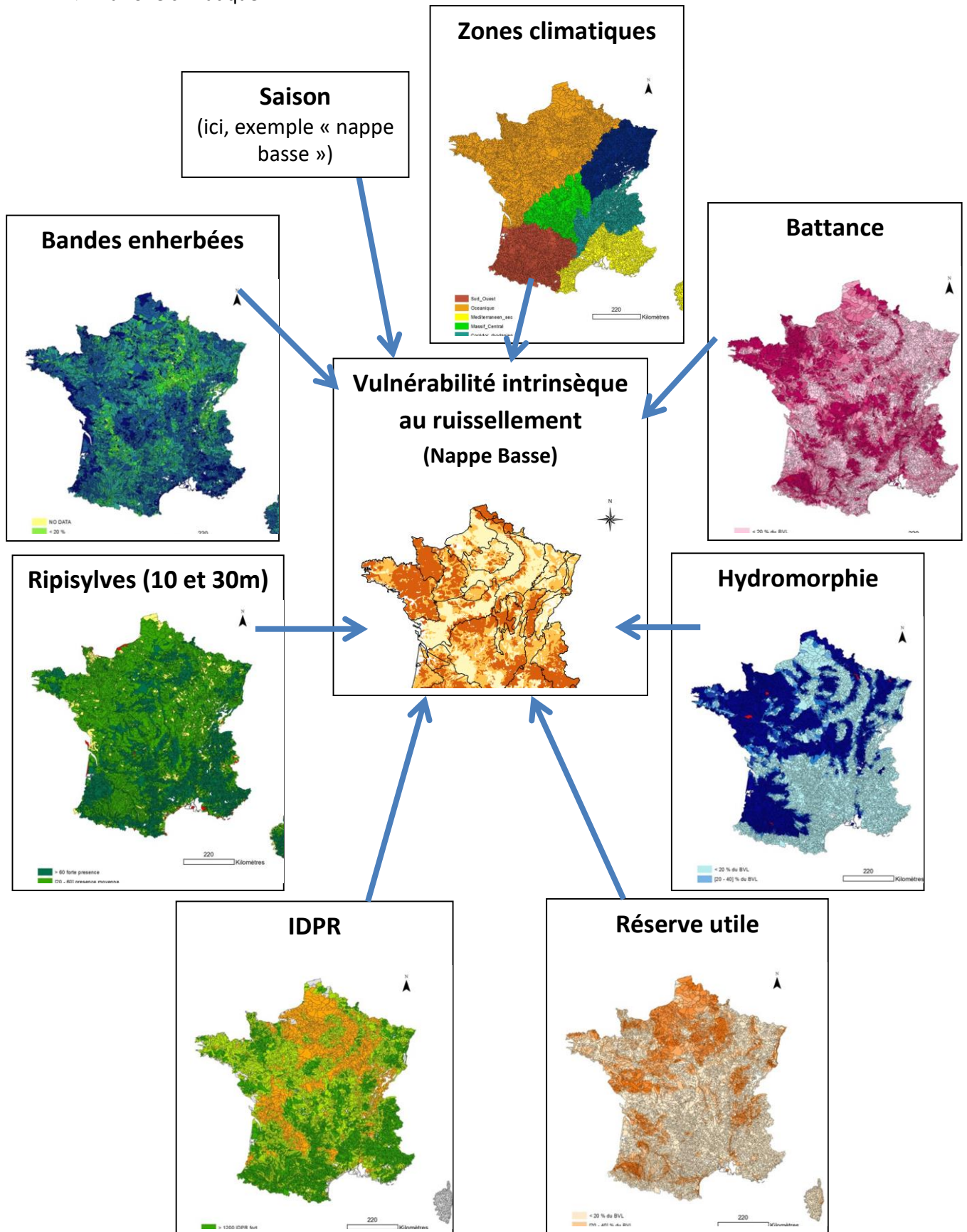


Figure 9 : Liens les cartes de variables d'entrée et la carte de vulnérabilité intrinsèque au ruissellement (exemple nappe basse)

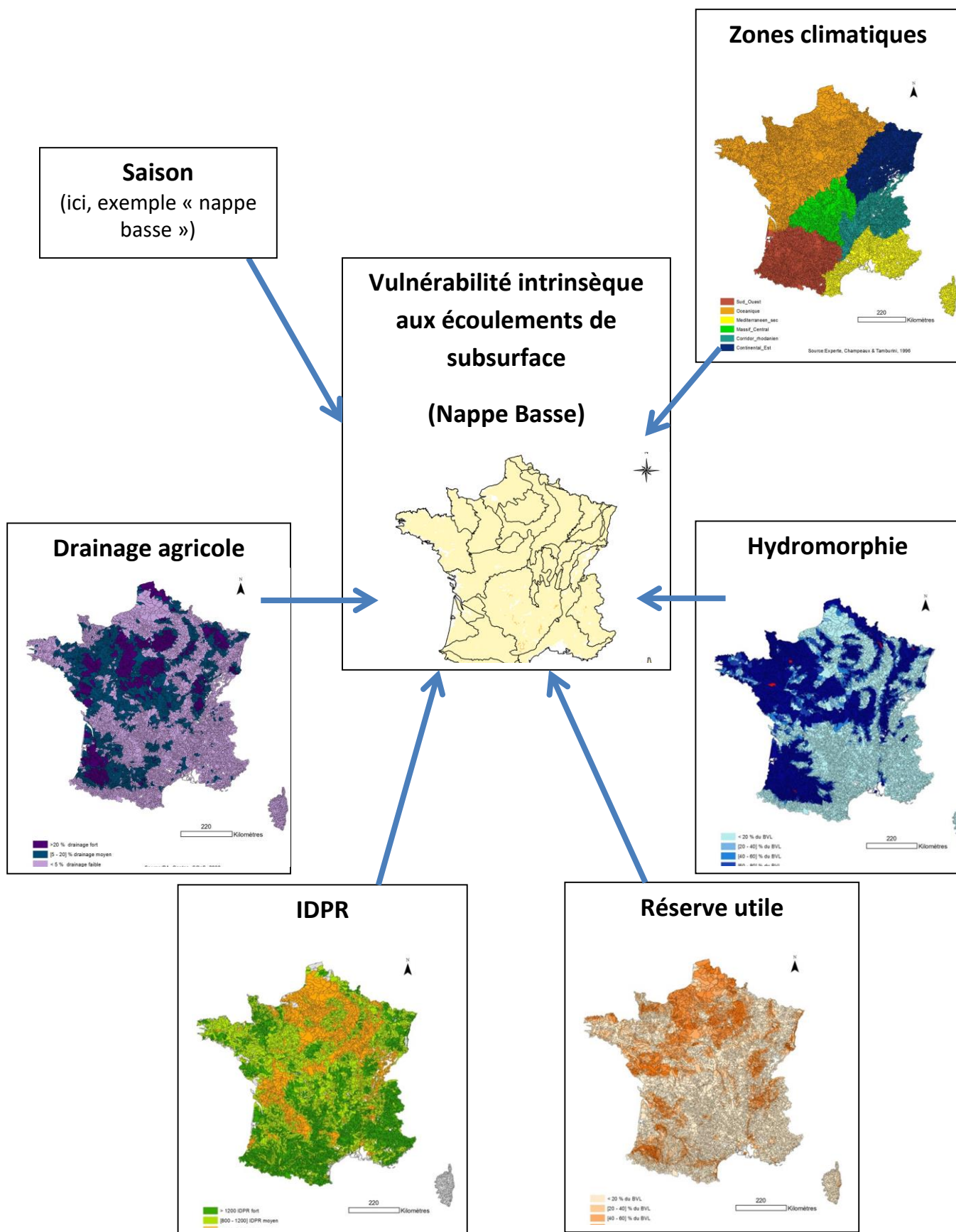


Figure 10 : Liens les cartes de variables d'entrée et la carte de vulnérabilité intrinsèque aux écoulements de subsurface (exemple nappe basse)

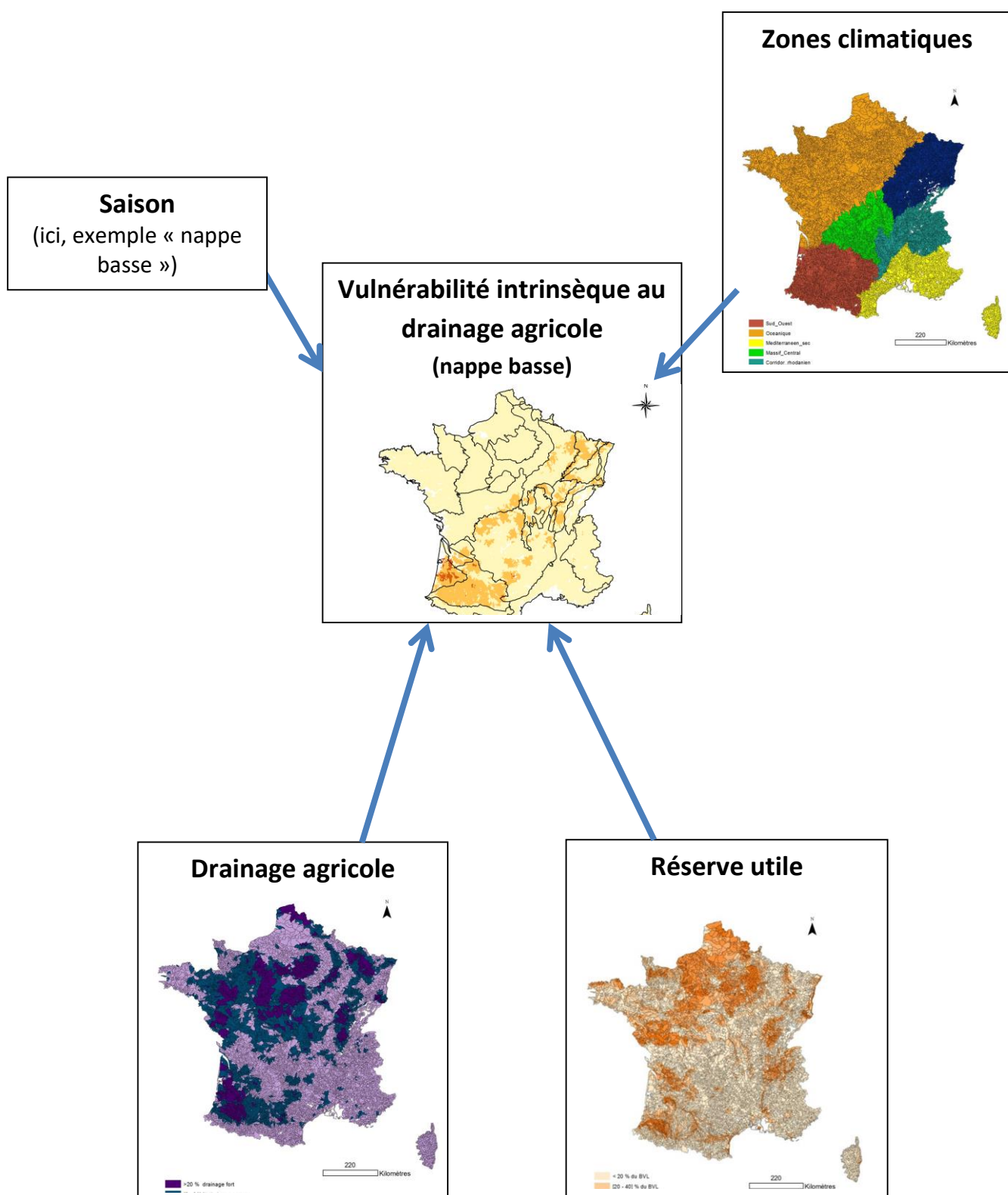


Figure 11 : Liens les cartes de variables d'entrée et la carte de vulnérabilité intrinsèque au drainage agricole (exemple nappe basse)

Vulnérabilité intrinsèque au drainage

On ne tient compte ici que du drainage enterré. La vulnérabilité est bien évidemment liée à la **présence de drains**, et est favorisée par une faible **réserve utile** et une forte **pluviométrie**.

Le drainage fonctionne en général en hiver, sauf dans le Sud-Ouest où il y en a également au printemps à cause de pluies intenses.

Les paramètres choisis pour estimer la vulnérabilité intrinsèque au drainage sont donc (voir Figure 11):

- ▶ Le pourcentage de la SAU qui est drainée
- ▶ La réserve utile
- ▶ La saison
- ▶ La zone climatique

Vulnérabilité spécifique via les transferts rapides

La vulnérabilité spécifique via les transferts rapides (nappe haute ou nappe basse) résulte du croisement entre les vulnérabilités intrinsèques au ruissellement et au drainage, et les propriétés des molécules. En effet, les transferts de produits phytosanitaires via l'eau dépendent non seulement des flux d'eau mais aussi des propriétés des substances. On ne prend pas en compte à ce stade les quantités en jeu.

Une vulnérabilité élevée au **drainage** et/ou au **ruissellement** entraîne une vulnérabilité spécifique élevée, à moduler par la teneur en matière organique du sol, le coefficient d'adsorption des substances sur le carbone organique (Koc), la solubilité des substances et leur demi-vie (DT50). En effet, si le sol est riche en **matière organique** et le **Koc** élevé, les substances se fixeront plus facilement dans le sol au lieu d'être entraînées dans l'eau. Par ailleurs, une **demi-vie** élevée induit une diminution plus lente de la quantité de produit appliqué au cours du temps et augmente donc la probabilité que les premiers transferts soient concentrés.

Enfin, une **fréquence de pluie** élevée augmente la probabilité qu'une pluie ait lieu peu après l'application de produit phytosanitaire et donc augmente le risque de contamination aiguë (avant que les substances n'aient eu le temps de se dégrader ou de se fixer).

Les paramètres pris en compte sont donc en résumé (voir Figure 12) :

- ▶ La vulnérabilité intrinsèque au drainage
- ▶ La vulnérabilité intrinsèque au ruissellement
- ▶ La fréquence de pluies (évaluées par expertise pour chaque zone climatique)
- ▶ Le Koc de la substance
- ▶ La demi-vie de la substance

Vulnérabilité spécifique via les transferts lents

De même que pour la vulnérabilité spécifique via les transferts lents, on tient compte de la vulnérabilité du milieu et des propriétés des molécules sans considérer, à ce stade, les quantités appliquées, et en tenant compte en plus des **flux de subsurface**.

Le coefficient d'adsorption est un facteur peu limitant pour la contamination chronique dans la mesure où des substances peu mobiles peuvent migrer à très faible dose au cours du temps si leur demi-vie est suffisamment importante. De ce fait, la **vitesse de dégradation** des substances est ici un paramètre considéré comme très important. Si elle est faible, la probabilité qu'il y ait une contamination du cours d'eau à long terme est assez élevée, même si sa vulnérabilité intrinsèque n'est

pas particulièrement importante. Le cumul de pluie a une influence sur les flux d'eau à long terme et donc sur la vulnérabilité chronique. Les paramètres pris en compte sont donc (voir Figure 13):

- ▶ La vulnérabilité intrinsèque au drainage
- ▶ La vulnérabilité intrinsèque au ruissellement
- ▶ La vulnérabilité intrinsèque aux écoulements de subsurface
- ▶ Le cumul de pluies
- ▶ La demi-vie des substances
- ▶ Le Koc des substances

Potentiel de contamination par les transferts rapides

Le potentiel de contamination par les transferts rapides d'une saison donnée dépend simplement de la vulnérabilité spécifique via les transferts rapides et de la pression d'usage phytosanitaire de cette même saison: plus l'une et/ou l'autre sont élevées, plus le potentiel de contamination est élevé.

Potentiel de contamination par les transferts lents

De même, le potentiel de contamination par les transferts lents dépend de la vulnérabilité spécifique via les transferts lents et de la pression d'usage phytosanitaire.

Cependant, dans ce cas, le lessivage lent des sols en nappe haute pouvant transférer des résidus de substances appliquées en nappe basse (bien qu'en faibles quantités) est pris en compte en rajoutant un lien entre la pressions d'usage nappe basse et le potentiel de contamination par les transferts lents en nappe haute. Ce lien n'a toutefois que peu d'effets sur le résultat final.

Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 » 3 : Liens entre variables

- La solubilité de la substance active, supprimée du modèle, était reliée à la vulnérabilité spécifique via les transferts rapides ;
- La densité de réseaux hydrographiques (nœud parent) était reliée à la vulnérabilité intrinsèque à la dérive avant d'être supprimée.

L'aléa érosif et la vulnérabilité intrinsèque à la dérive étaient reliées à la vulnérabilité spécifique via les transferts rapides avant d'être de même supprimés du modèle.

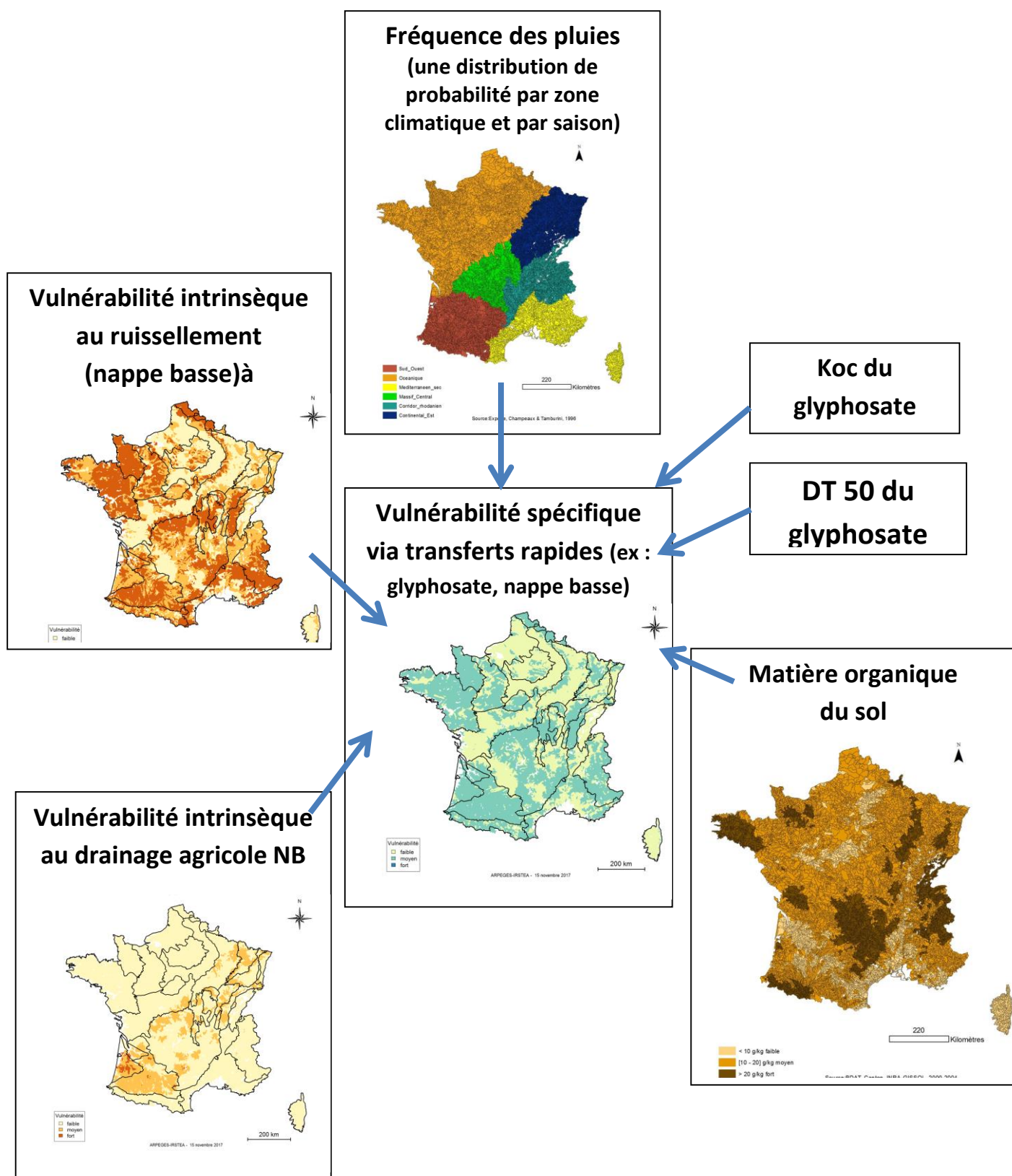


Figure 12 : Liens entre la carte de vulnérabilité spécifique via les transferts rapides et les cartes de ses nœuds parents (exemple glyphosate nappe basse)

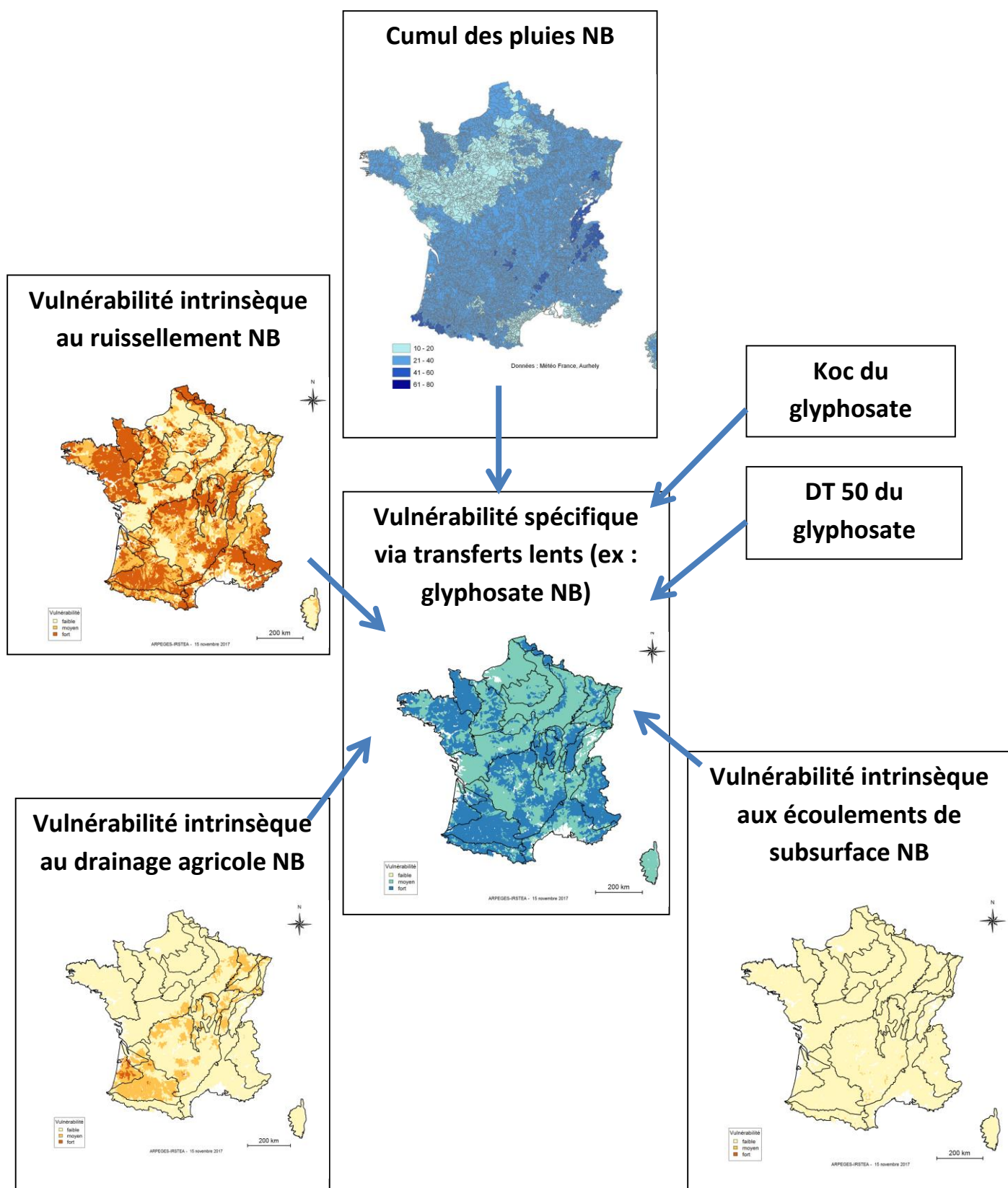


Figure 13 : Liens entre la carte de vulnérabilité spécifique via les transferts lents et les cartes de ses nœuds parents (exemple glyphosate nappe basse)

3.3.2. Construction de la structure probabiliste

La construction de la structure probabiliste se résume en une seule et 4^{ème} étape de la construction globale du réseau bayésien. Il s'agit de compléter une table de probabilité conditionnelle pour chaque nœud enfant du réseau. Toutes les tables ont été renseignées par expertise.

L'explication du principe des tables de probabilités est à retrouver dans l'encadré « Les fondements mathématiques 2 ». On comprend notamment que l'ajout d'une nouvelle variable démultiplie le nombre de lignes à renseigner dans la table directement impactée. Dans une moindre mesure, l'ajout de nouveaux états d'une variable nécessite de renseigner un nombre assez conséquent de nouvelles lignes.

Sur								la		
nappe haute	Océanique	faible	moyen	faible	faible	faible	absence	20	50	30
nappe haute	Océanique	faible	moyen	faible	faible	faible	présence	30	50	20
nappe haute	Océanique	faible	moyen	faible	faible	moyen	absence	30	50	20
nappe haute	Océanique	faible	moyen	faible	faible	moyen	présence	40	50	10
nappe haute	Océanique	faible	moyen	faible	faible	fort	absence	40	50	10
nappe haute	Océanique	faible	moyen	faible	faible	fort	présence	60	40	0
nappe haute	Océanique	faible	moyen	faible	moyen	faible	absence	20	50	30
nappe haute	Océanique	faible	moyen	faible	moyen	faible	présence	30	50	20
nappe haute	Océanique	faible	moyen	faible	moyen	moyen	absence	30	50	20
nappe haute	Océanique	faible	moyen	faible	moyen	moyen	présence	40	50	10
nappe haute	Océanique	faible	moyen	faible	moyen	fort	absence	40	50	10
nappe haute	Océanique	faible	moyen	faible	moyen	fort	présence	60	40	0
nappe haute	Océanique	faible	moyen	faible	fort	faible	absence	20	50	30
nappe haute	Océanique	faible	moyen	faible	fort	faible	présence	30	50	20
nappe haute	Océanique	faible	moyen	faible	fort	moyen	absence	30	50	20
nappe haute	Océanique	faible	moyen	faible	fort	moyen	présence	40	50	10
nappe haute	Océanique	faible	moyen	faible	fort	fort	absence	40	50	10
nappe haute	Océanique	faible	moyen	faible	fort	fort	présence	60	40	0
nappe haute	Océanique	faible	moyen	moyen	faible	faible	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	faible	présence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	moyen	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	moyen	présence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	fort	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	fort	présence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	moyen	faible	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	moyen	faible	présence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	moyen	moyen	absence	33.333	33.333	33.333

Figure 14 ci-dessous, on trouve un exemple concret de table de probabilité, extrait du réseau bayésien d'ARPEGES, pour la vulnérabilité intrinsèque au ruissellement. Une ligne correspond à une combinaison d'états des 8 nœuds parents sur les 5832 combinaisons possibles⁶. Il s'agit plutôt de deux courts extraits de la même table (comme le reflète la position du curseur de déroulement sur la droite de chacun des extraits), l'un pour la modalité « nappe basse » de la variable « saison » l'autre pour la modalité « nappe haute », les combinaisons pour les autres variables étant les mêmes entre les deux extraits. Cela permet de visualiser l'influence de la saison sur ces quelques combinaisons.

Sur la Figure 15, qui correspond au réseau bayésien final après renseignement de toutes les tables de probabilités, on constate l'effet de ces tables sur les probabilités des états de chaque variable en

⁶ La table qui contient les probabilités conditionnelles pour le nœud « vulnérabilité intrinsèque au ruissellement » est constituée de 5832 lignes si on considère 6 zones climatiques (5 nœuds parents qui peuvent prendre 3 valeurs, 2 nœuds parents qui peuvent prendre 2 valeurs et 1 nœud parent, les zones climatiques, qui peut prendre 6 valeurs, soit $3^5 \times 2^2 \times 6$ croisements possibles). Le nombre de lignes doublerait si on considérait 12 zones.

comparant ce réseau à celui de la Figure 8 : ceux-ci ne sont plus équiprobables, mais sont la résultante de la formule des probabilités décrite dans l'encadré « **Les fondements mathématiques n°3** » de la **partie 3.5**. Seuls les états des variables d'entrée sont équiprobables car non encore renseignés.

Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 » : Tables de probabilités

- Table de la vulnérabilité spécifique via les transferts rapides :

Du fait de la suppression de l'érosion comme variable d'entrée, une troncature de la table de probabilités existant au préalable a été réalisée en ne gardant que la partie de la table où l'érosion était forte, en réajustant à la marge les probabilités conditionnelles pour rééquilibrer les distributions de probabilités.

- Tables de probabilités des potentiels de contamination :

Le nombre d'états des potentiels de contamination est passé de 3 à 5, de même que le nombre d'états des pressions d'usage phytosanitaire. Par voie de conséquence, la configuration de la table de probabilités des potentiels de contamination a changé, et les probabilités contenues dans les tables précédentes ont été modifiées et complétées pour les nouvelles lignes induites par les nouveaux états. Ce travail a été réalisé par expertise.

Les possibilités de l'interface R : Tables de probabilités

Il n'y a pas d'action possible pour l'utilisateur sur cette partie « Tables de probabilités », car cela engendrerait des modifications majeures sur le modèle en lui-même.

nappe haute	Océanique	faible	moyen	faible	faible	faible	absence	20	50	30
nappe haute	Océanique	faible	moyen	faible	faible	faible	presence	30	50	20
nappe haute	Océanique	faible	moyen	faible	faible	moyen	absence	30	50	20
nappe haute	Océanique	faible	moyen	faible	faible	moyen	presence	40	50	10
nappe haute	Océanique	faible	moyen	faible	faible	fort	absence	40	50	10
nappe haute	Océanique	faible	moyen	faible	faible	fort	presence	60	40	0
nappe haute	Océanique	faible	moyen	faible	moyen	faible	absence	20	50	30
nappe haute	Océanique	faible	moyen	faible	moyen	faible	presence	30	50	20
nappe haute	Océanique	faible	moyen	faible	moyen	moyen	absence	30	50	20
nappe haute	Océanique	faible	moyen	faible	moyen	moyen	presence	40	50	10
nappe haute	Océanique	faible	moyen	faible	moyen	fort	absence	40	50	10
nappe haute	Océanique	faible	moyen	faible	moyen	fort	presence	60	40	0
nappe haute	Océanique	faible	moyen	faible	fort	faible	absence	20	50	30
nappe haute	Océanique	faible	moyen	faible	fort	faible	presence	30	50	20
nappe haute	Océanique	faible	moyen	faible	fort	moyen	absence	30	50	20
nappe haute	Océanique	faible	moyen	faible	fort	moyen	presence	40	50	10
nappe haute	Océanique	faible	moyen	faible	fort	fort	absence	40	50	10
nappe haute	Océanique	faible	moyen	faible	fort	fort	presence	60	40	0
nappe haute	Océanique	faible	moyen	moyen	faible	faible	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	faible	presence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	moyen	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	moyen	presence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	fort	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	faible	fort	presence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	moyen	faible	absence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	moyen	faible	presence	33.333	33.333	33.333
nappe haute	Océanique	faible	moyen	moyen	moyen	moyen	absence	33.333	33.333	33.333

Figure 14 : Extrait de la table de probabilités du noeud Vulnérabilité intrinsèque au ruissellement

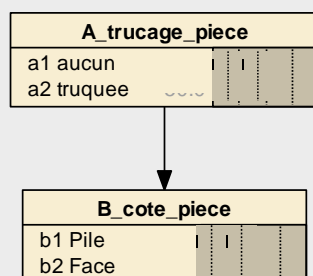
Les fondements mathématiques 2 : Structure probabiliste des réseaux bayésiens

- Notion de probabilités conditionnelles

Pour chaque **combinaison d'états** des nœuds parents, il faut donner une probabilité conditionnelle à chacun des **états (=événements)** possibles du nœud enfant considéré. Chaque combinaison d'états des nœuds parents peut être vue comme une **expérience** et les probabilités conditionnelles comme le **résultat moyen d'un grand nombre de répétitions de cette expérience**.

Par exemple, prenons une expérience de Pile ou Face :

- En l'absence d'information, je m'attends à avoir 50% de chances de faire Pile, et 50% de chances de faire Face : sur une expérience de 100 lancers, je réalise en moyenne 50 Piles et 50 Faces (mais l'expérience peut aussi donner 60 Piles et 40 Faces, ou encore le cas extrême de 100 Faces pour 0 Pile, mais la probabilité que ces événements se réalisent est plus faible). Il n'y a dans ce cas aucune probabilité conditionnelle.
- A présent, considérons que l'on m'informe que la pièce est peut-être truquée vers les Piles. Je dois considérer les deux cas : si la pièce n'est pas truquée, j'ai toujours une répartition de probabilités de 50% - 50% ; si la pièce est truquée, j'augmente la probabilité des Piles, mettons par exemple 70% - 30%.
 - ⇒ Cela ne m'assure pas que la pièce est truquée ou pas. Je n'en suis pas sûr, je considère donc les deux cas possibles et je réalise un pari sur le résultat dans les deux cas.
 - ⇒ Que la pièce soit truquée ne m'assure pas que je fasse un Pile. J'ai simplement plus de chance à chaque tirage d'en faire un (la probabilité de faire Pile, conditionnellement au fait que la pièce soit truquée, a augmenté), mais il arrive régulièrement qu'un tirage donne un Face.



Dans cet exemple très simple, le nœud enfant ne dépend ici que d'un seul nœud parent (pas de combinaison d'états). On a dans ce cas un nœud parent (trucage de la pièce) avec deux états et expériences (aucun/truquée) et un nœud enfant (résultat moyen des lancers) avec deux événements possibles (Pile / Face).

Illustrations à partir du logiciel NETICA

On notera $p(B|A)$ (on lit probabilité de B sachant A) la probabilité de la variable aléatoire B (dans notre exemple : côté de la pièce) sachant la variable A (trucage de la pièce).

Si un nœud A pouvant prendre les états/événements $a1$ et $a2$ est conditionnellement lié à un nœud B pouvant prendre les états $b1$ et $b2$, alors la probabilité que B prenne la valeur $b1$ sachant que A prend la valeur $a2$ sera notée $p(B=b1 | A=a2)$.

Dans l'exemple précédent : $p(B=Pile | A = truquée) = 70\%$.

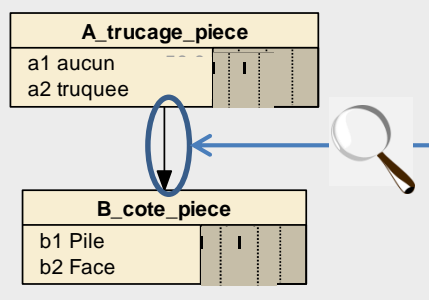
Les fondements mathématiques 2: Structure probabiliste des réseaux bayésiens (suite)

- Table de probabilité conditionnelle

Derrière chaque convergence de flèches de la structure graphique, on décrit le lien de causalité par une distribution de probabilités conditionnelles. Cette distribution décrit, pour chaque combinaison d'états des nœuds parents (= chaque expérience), les probabilités des états possibles d'un nœud enfant, conditionnellement à ses nœuds parents, c'est-à-dire connaissant, même hypothétiquement, l'état pris par chacun de ses nœuds parents.

On résume cette distribution conditionnelle à l'aide d'une table de probabilités conditionnelles, dans laquelle chaque ligne est une combinaison d'états des nœuds parents et les probabilités associées des états du nœud enfant.

Si l'on reprend l'exemple très simple de l'encadré précédent, on obtient cette table de probabilités conditionnelles, associée à l'unique flèche :



A_trucage_piece

a1 aucun
a2 truquee

B_cote_piece

b1 Pile
b2 Face

B_cote_piece Table (in Bayes net exemple_trucage_piece)

Node: **B_cote_piece** Apply Reset

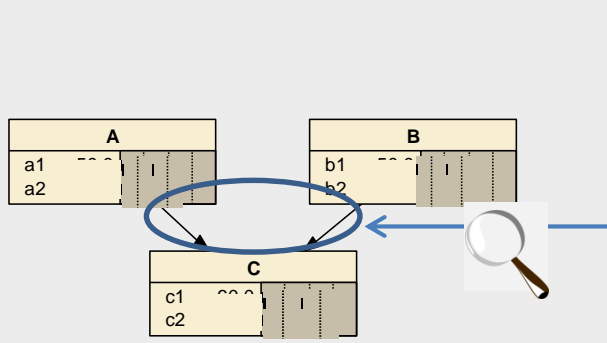
Chance ▼ % Probability ▼

A_trucage_piece	b1 Pile	b2 Face
a1 aucun	50	50
a2 truquee	70	30

On retrouve, entourée sur la deuxième ligne, la probabilité conditionnelle citée en exemple dans l'encadré précédent : $p(B=Pile \mid A = \text{truquée}) = 70\%$.

Cependant, les nœuds enfants d'un réseau bayésien sont en général issus du croisement d'au moins deux nœuds parents.

L'exemple suivant présente un cas théorique croisant deux variables :



A

a1
a2

B

b1
b2

C

c1
c2

C Table (in Bayes net exemple_reseau_3noeuds)

Node: **C** Apply Reset

Chance ▼ % Probability ▼

A	B	c1	c2
a1	b1	40	60
a1	b2	80	20
a2	b1	50	50
a2	b2	70	30

Par exemple, sur la première ligne, on lira que si la variable A prend la valeur a1 et que la variable B prend la valeur b1, alors la variable C a 40% de chances de prendre la valeur c1 et 60% de chances de prendre la valeur c2. On notera donc :

- $p((C=c1 \mid (A=a1 \cap B=b1))) = 0,4$
- $p((C=c2 \mid (A=a1 \cap B=b1))) = 0,6$

Les deux nœuds parents et le nœud enfant ayant tous, dans cet exemple, deux états potentiels, la table de probabilité comporte 4 lignes (= 4 combinaisons des états des nœuds parents), 2 colonnes et 8 probabilités conditionnelles à renseigner.

Il apparaît que plus le nombre de nœuds parents et/ou plus le nombre d'états possibles pour les nœuds, tant parents qu'enfants, augmentent, plus les tables de probabilités conditionnelles deviennent grandes et donc compliquées à renseigner. Soit y le nombre de nœuds parents, (n_1, n_2, \dots, n_y) le nombre d'états respectif de chaque nœud parent, et n_e le nombre d'états du nœud enfant, alors le nombre de lignes de la table est $n_1 \times n_2 \times \dots \times n_y$, et n_e le nombre de colonnes. Par exemple, si tous les nœuds parents conditionnant un nœud enfant se décrivent selon 3 états, alors le nombre de lignes sera de 3^y .

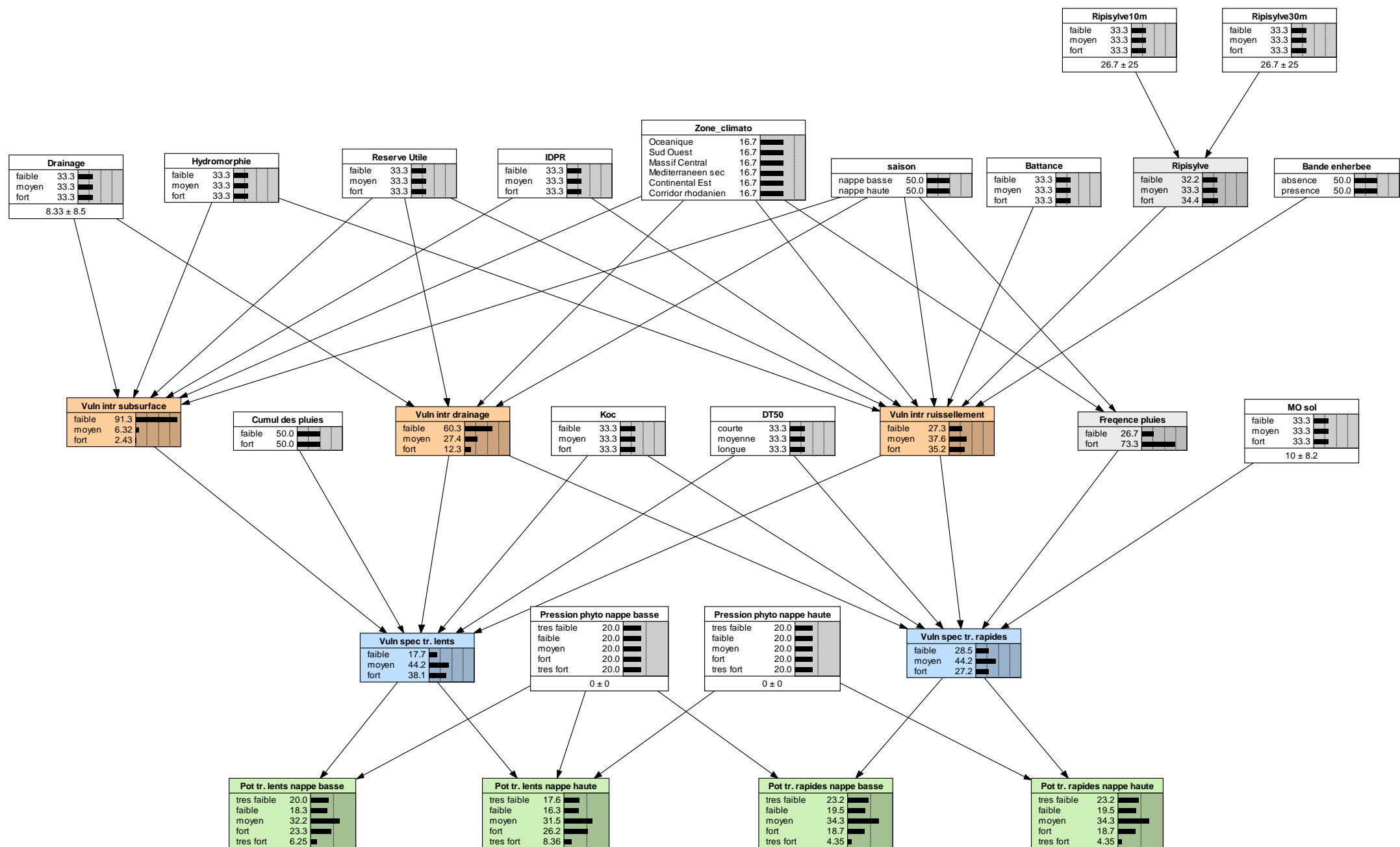


Figure 15 : Construction de la structure probabiliste d'ARPEGES

3.4. Données d'entrée

Pour renseigner les données d'entrée du réseau bayésien d'ARPEGES, 4 étapes sont nécessaires : collecte des données, ajustement à la maille de calcul, discrétisation (si la donnée est continue) et traduction en n-uplets de probabilités. Les deux premières étapes se font en-dehors de l'interface R, les deux dernières sont prises en charge par l'outil. Les paragraphes de cette partie détaillent chacune de ces étapes.

3.4.1. Préparation des données d'entrée

La préparation des données d'entrée recouvre la collecte des données et l'ajustement à la maille de calcul. Le premier paragraphe présente une synthèse de ces deux étapes, puis celles-ci sont détaillées dans les paragraphes suivants.

3.4.1.1. Synthèse des données d'entrée

Le Tableau 4 rassemble les informations synthétiques concernant les données d'entrée utilisées pour le projet « ARPEGES EDL 2019 ».

3.4.1.2. Collecte des données d'entrée : description et limites

Une donnée par variable est intégrée dans le modèle ARPEGES. Chaque donnée est décrite ci-dessous (année de production de la donnée, échelle, ...), avec un rappel des nœuds enfants conditionnés par chaque donnée.

Données pédologiques

○ IDPR :

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque au ruissellement**
- ⇒ **vulnérabilité intrinsèque aux écoulements de subsurface**

L'IDPR résulte d'une comparaison entre le réseau hydrographique théorique (calculé à partir du MNT) et le réseau réel. Il est calculé pour des mailles de 1 km². Si le réseau réel est moins dense que le réseau théorique, cela traduit le fait que l'infiltration de l'eau est majoritaire, tandis que s'il est plus dense les flux se font majoritairement à destination des eaux de surface (ESU).

L'IDPR permet de prédire le milieu destinataire de l'eau (ESU vs ESO), mais n'indique pas les voies spécifiques de transfert (ruissellement, drainage...). Par exemple, si l'eau va majoritairement vers les ESU, les flux peuvent être du ruissellement par saturation, du ruissellement hortonien, du drainage, des flux de subsurface... De même, l'eau peut d'abord ruisseler avant de s'infiltrer ou s'infiltrer avant d'alimenter une rivière plus en aval (milieux karstiques par exemple).

L'IDPR permet de prédire la destination de l'eau, mais n'indique pas à lui seul les voies spécifiques de transfert. Pour cela, il doit s'accompagner des informations pédologiques (battance, réserve utile ...).

L'IDPR utilisé date de 2010, mais une mise à jour de cette donnée par le BRGM est à venir prochainement.

Variables	Données			Méthode d'évaluation au niveau du BVME	Fiche synthétique
	Nom	Unité spatiale ou échelle spatiale	Date		
Ruissellement/infiltration	IDPR	1/50 000	2010	Agrégation par la moyenne (cas n°1)	ANNEXE 2
Réserve utile	BDGSF	UCS	1998	Agrégation probabiliste (cas n°2)	ANNEXE 3
Drainage	RGA	Canton	2000	Désagrégation-réagrégation (cas n°2)	ANNEXE 4
Hydromorphie	Carte Lagacherie	PRN	1987	Agrégation probabiliste (cas n°2)	ANNEXE 5
Battance	BDGSF	UCS	2002	Agrégation probabiliste (cas n°2)	ANNEXE 6
Ripisylve 30 m	BD TOPO®	1/ 25 000	2010	Agrégation par la somme (cas n°1)	ANNEXE 7
Ripisylve 10 m	BD TOPO®	1/ 25 000	2010	Agrégation par la somme (cas n°1)	ANNEXE 8
Bande enherbée	BD TOPO®	1/ 25 000	2010	Agrégation par la somme (cas n°1)	ANNEXE 9
Zone climatique	Délimitation experte	-	-	Désagrégation par Jointure spatiale (cas n°3)	ANNEXE 10
Cumuls de pluies	AURELHY	1 km²	1971-2000	Agrégation par la moyenne (cas n°1)	ANNEXE 11
Matière organique du sol	BDAT	Cantons	2000-2004	Agrégation probabiliste (cas n°2)	ANNEXE 12
DT 50	SIRIS	-	2012	-	
KOC	SIRIS	-	2012	-	
Pression d'usage phytosanitaire nappe haute et nappe basse	BNV-d spatialisée + Coefficients de saisonnalité estimés par connaissance experte	Communes	2015	Agrégation par la somme (cas n°1)	ANNEXE 13

Tableau 4 : Synthèse des données d'entrée d'ARPEGES

En marron les paramètres d'entrée pour évaluer une ou des vulnérabilités intrinsèques, en vert pour une ou des vulnérabilités spécifiques, et en rose pour la pression d'usage phytosanitaire (UCS : Unité Cartographique des Sols de la BD cartographique des sols ; PRN : Petite région naturelle ; HER : Hydro-ÉcoRégion niveau 1)

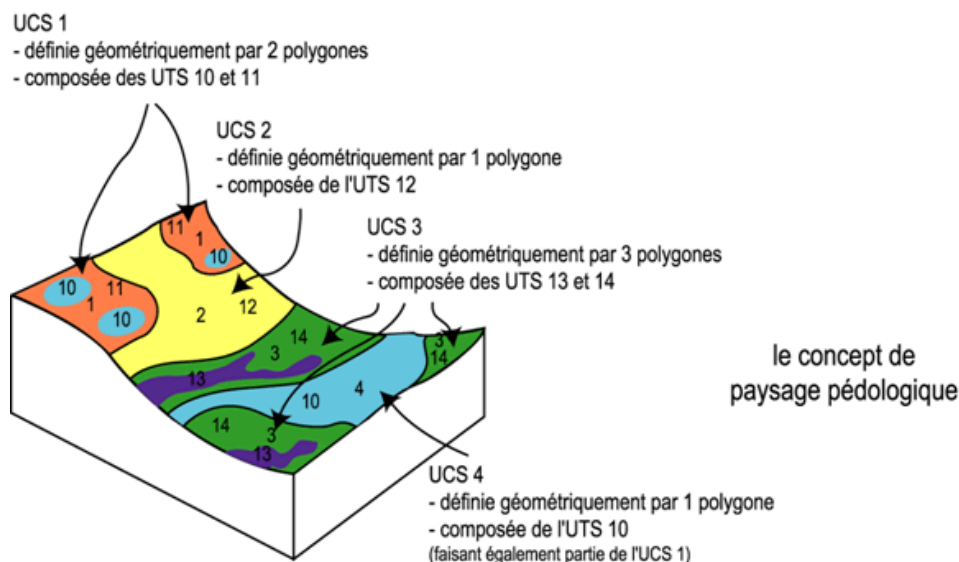
○ Battance :

Nœuds enfants :

⇒ vulnérabilité intrinsèque au ruissellement

La battance introduite dans ARPEGES provient de la BDGSF (Base de Données Géographique des Sols de France au 1/1.000.000^{ème}, 1988) du GIS Sol (Groupement d'Intérêt Scientifique Sol). La BDGSF est composée tout d'abord d'une carte des Unités Cartographiques de Sols (UCS). Ces unités ne sont pas homogènes, mais sont constituées d'Unités Typologiques de Sols (UTS) qui ne sont pas cartographiables à cette échelle (Figure 16). A chaque UCS correspond donc une liste d'UTS avec la proportion de l'UCS qu'elles représentent. Chaque UTS est décrite par une quarantaine de variables.

Ces variables peuvent être des observations du sol ou des données obtenues par des règles de pédotransfert, comme la sensibilité à la formation d'une croûte de battance (obtenue à partir de la texture et du type de sol) (Le Bissonnais et al. 2002).



© INRA

Figure 16 : Concepts d'UCS et d'UTS (INRA)

La particularité de cette source de données est le fait que chaque variable est discrète : la sensibilité à la formation d'une croûte de battance est par exemple découpée en 5 modalités dans la BDGSF, et doit être découpée en 3 modalités pour ARPEGES (cf §3.3.1.2). Les modalités « très faible » et « faible » ont donc été regroupées en « faible », ainsi que les modalités « fort » et « très fort » en « fort ».

Une carte pour chaque modalité résulte de cette méthode.

Ce regroupement a eu pour conséquence de changer quelque peu la spatialisation de la donnée battance. Par exemple, avant regroupement, le Pays de Caux (Normandie) a des surfaces classées en « très fort », tandis que dans la région de Saint-Malo, la plus haute modalité n'est que la modalité « fort ». Après regroupement, la spatialisation est inversée (plus grande surface de « fort » dans la région de Saint-Malo qu'en Pays de Caux), sans doute car la surface en « fort » dans les alentours de

Saint-Malo est plus grande que la surface en « fort » et « très fort » réunis en Pays de Caux (voir Figure 17 pour illustration et Annexe 6 pour plus de précisions).

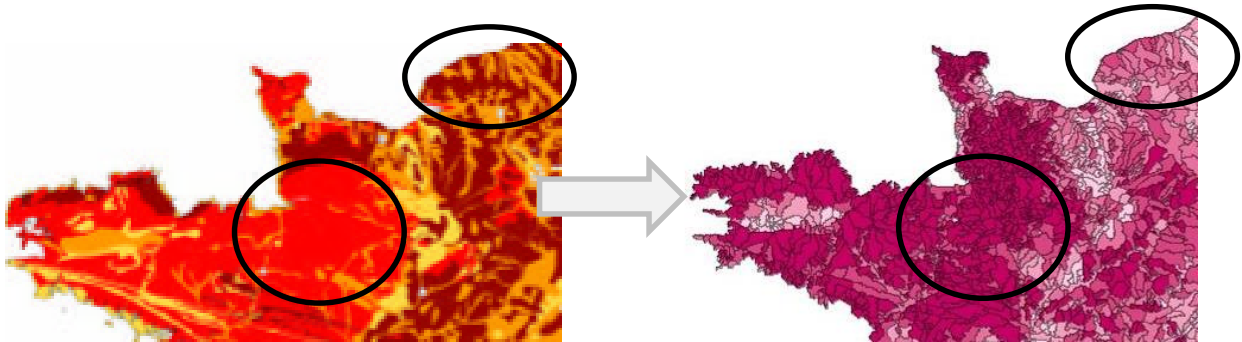


Figure 17 : Comparaison des cartes avant (à gauche) et après (à droite) discrétisation de la donnée battance pour la zone Bretagne-Normandie

○ Réserve utile :

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque au ruissellement**
- ⇒ **vulnérabilité intrinsèque aux écoulements de subsurface**
- ⇒ **vulnérabilité intrinsèque au drainage agricole**

La source de données de la réserve utile est la même que celle de la battance, à savoir la BDGSF. De même, un regroupement de modalités a été opéré : les modalités « très faible » (< 50 mm) et « faible » (50-100 mm) ont été regroupées en « faible », et les modalités « fort » (150-200 mm) et « très fort » (≥ 200 mm) en « fort ».

La limite inhérente au regroupement de modalités existe aussi pour la réserve utile, même si l'on n'a pas constaté de changements de spatialisation induits.

○ Carbone organique du sol :

Nœuds enfants :

- ⇒ **vulnérabilité spécifique aux transferts rapides**

La donnée « carbone organique du sol » provient de la BDAT (Base de Données des Analyses de Terre) du GIS Sol. Cette base de données recense les résultats des analyses de terre. Ces analyses sont réalisées généralement à la demande des agriculteurs pour gérer au mieux la fertilisation. Sont disponibles la texture, le pH, la capacité d'échange cationique et les teneurs en nitrates, en phosphore et en carbone organique. Ces données sont accessibles par des statistiques par canton (médiane, moyenne...). Les données utilisées concernent la période 2000-2004.

○ Hydromorphie :

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque au ruissellement**
- ⇒ **vulnérabilité intrinsèque aux écoulements de subsurface**

La carte d'hydromorphie utilisée pour ARPEGES est la carte de l'hydromorphie à l'échelle des Petites Régions Naturelles (PRN) produite par Lagacherie (1987). Cette carte a été élaborée à partir d'une enquête auprès des pédologues locaux et d'une étude des cartes pédologiques existantes. Cinq catégories de petites régions naturelles ont été définies.

Pour cette source de données aussi, la donnée est discrète (déjà classée en 5 modalités) et exprimée en pourcentage de sols hydromorphes dans la petite région naturelle. De même que pour la battance et la réserve utile, des modalités ont été regroupées pour n'en avoir plus que 3 :

- Faible dans ARPEGES = très faible (moins de 20 % de sols hydromorphes) et faible (moins de 20% de sols hydromorphes – engorgement ponctuel)
- Fort dans ARPEGES = fort (20-50% de sols hydromorphes) + très fort (50% de sols hydromorphes).

Données sur les aménagements des sols et du paysage :

○ **Drainage agricole :**

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque aux écoulements de subsurface**
- ⇒ **vulnérabilité intrinsèque au drainage agricole**

La donnée « drainage agricole » retenue provient du recensement agricole (RA) Agreste, sous la forme « SAU drainée ». Il s'agit d'un recensement de l'ensemble des exploitations françaises. Les variables sont collectées au niveau communal. Cependant, elles ne sont pas accessibles dans toutes les communes à cause du secret statistique. Nous les considérons donc à l'échelle des cantons. Bien que le RA 2010 soit aujourd'hui disponible, c'est le RA 2000 qui est toujours utilisé dans ARPEGES, car la variable « drainage agricole » aurait perdu en qualité (Comm. pers. Agence de l'Eau Loire-Bretagne).

○ **Ripisylves 10m et 30m :**

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque au ruissellement**

Les ripisylves 10m et 30m sont disponibles dans la BD TOPO® de l'IGN (Institut National Géographique, 2010). Cette base décrit des éléments du paysage sous forme de vecteurs avec une précision métrique, classés selon une thématique adaptée.

Les données « ripisylves 10m et 30m » n'existent pas à proprement parler dans la BD TOPO®. Pour les obtenir, le pourcentage surfacique des polygones « forêt » de la couche de végétation découpés dans un buffer de 30 m autour du réseau hydrographique du bassin versant a été calculé (réseau BDTOPO issu du réseau SYRAH (Valette et al, 2008)).

Cette méthode ne permet pas d'évaluer si la forêt est bien répartie le long du cours d'eau ou pas, ni si elle est située entre les zones de culture et le cours d'eau, ce qui a des conséquences différentes sur les transferts de produits phytosanitaires.

- **Bandes enherbées :**

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque au ruissellement**

La source de données pour les bandes enherbées est également la BD TOPO®. Faut de données sur la présence réelle de bandes enherbées en France, nous avons considéré le pourcentage de linéaire de cours d'eau qui doit légalement être protégé par des bandes enherbées, c'est-à-dire les cours d'eau qui portent un nom sur les cartes IGN au 1/25.000ème.

Cette donnée sera sans doute à actualiser compte tenu de la révision du classement des cours d'eau, toujours en cours en 2018.

Données climatiques :

- **Zones climatiques :**

Nœuds enfants :

- ⇒ **vulnérabilité intrinsèque au ruissellement**
- ⇒ **vulnérabilité intrinsèque aux écoulements de subsurface**
- ⇒ **vulnérabilité intrinsèque au drainage agricole**

Les zones climatiques ont été définies à l'aide d'un zonage existant basé sur les séries de précipitations (Champeaux & Tamburini 1996). Ce zonage était initialement composé de 12 zones, ce qui était trop compte tenu de notre méthode de croisement des données (car cela aurait conduit à un trop grand nombre de combinaisons possibles). Nous avons donc fusionné des zones pour n'en garder que 6. Notamment, les zones montagneuses ont été fusionnées à la zone non montagneuse la plus proche, même si elles présentent une pluviométrie plus élevée, car ce ne sont pas les territoires qui vont être les plus déterminants dans la contamination par les produits phytosanitaires du fait des usages moindres.

Les zones ainsi obtenues correspondent, à dire d'experts, à des zones globalement homogènes du point de vue de la fréquence et de l'intensité des pluies.

Les cumuls de pluie étant aussi variables à l'intérieur des zones climatiques qu'entre zones (notamment à cause du relief), nous avons préféré introduire ce paramètre spécifique déterminé grâce à la méthode Aurelhy (Analyse Utilisant le RELief pour l'Hydrométéorologie) de Météo France (Canellas et al, 2014).

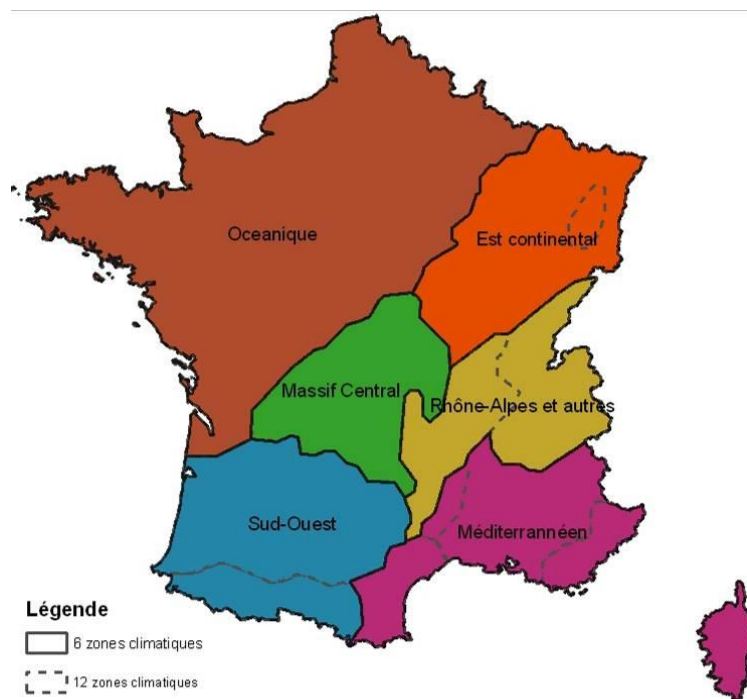


Figure 18 : Carte des 6 zones climatiques définies pour la méthode

○ Cumul des pluies :

Les cumuls de pluies proviennent de la méthode Aurelhy. Il s'agit d'une méthode d'interpolation basée sur la liaison statistique existant entre les précipitations et le relief environnant. Plusieurs données sont disponibles pour des mailles de 1 km² : normales de cumul de précipitations, mensuelles ou annuelles. Ce sont les normales mensuelles sur la période 1971-2000 qui ont été utilisées.

Caractéristiques des substances actives

Les propriétés des substances actives des produits phytosanitaires sont évaluées en vue de leur mise sur le marché. Plusieurs bases de données existent reprenant ces données complétées éventuellement par d'autres sources (Agritox, Footprint-PPDB et SIRIS-pesticides). Parmi celles-ci, nous avons sélectionné la base SIRIS-pesticides car elle est directement accessible sous forme de table gratuitement, et fournit une seule valeur par caractéristique de chaque substance active.

L'outil SIRIS-Pesticides⁷ de l'Ineris permet d'évaluer quelles substances il faut rechercher en priorité dans les eaux. Il se base sur les quantités de produits appliquées et sur les propriétés des substances actives. Elle contient notamment la DT50 (demi-vie) et le Koc (coefficient d'absorption sur le carbone organique). La dernière mise à jour de cette table date de 2012.

⁷ <http://www.ineris.fr/siris-pesticides/>

- Pression d'usage phytosanitaire :

⇒ **Donnée pour la pression d'usage annuelle**

Description générale de la donnée

On rappelle au lecteur une question de vocabulaire : le terme « **pression** » n'est employé dans ce guide que pour signifier la **pression liée aux usages phytosanitaires**, c'est-à-dire la **quantité de substance active appliquée par unité de surface (il s'agit donc d'une donnée en g/ha)**. Il s'agit d'une pression potentielle, avant transfert vers le cours d'eau. Il ne s'agit en aucun cas de la pression exercée sur les masses d'eau, qui constitue le résultat du transfert des substances actives vers les cours d'eau.

La caractérisation de l'usage des produits phytosanitaires est difficile, tant du point de vue des substances utilisées selon les cultures et les zones considérées que des quantités appliquées. Le caractère lacunaire des données d'usage constitue souvent un facteur limitant des études du potentiel de contamination. Pour l'application d'ARPEGES de surcroît, il est préférable de disposer de données acquises de manière homogène pour tout le territoire. Dans le cadre de cette étude, le choix a été fait de s'appuyer sur les données de la Banque Nationale des Ventes de produits phytopharmaceutiques par les Distributeurs agréés (BNVD).

La BNVD est une base qui existe depuis 2008. Les distributeurs doivent déclarer les volumes de chaque produit qu'ils vendent, en vue du paiement de la redevance pour pollutions diffuses définie dans le cadre de la loi sur l'eau et les milieux aquatiques (LEMA) de décembre 2006. Initialement déclarées à la commune de chaque point de vente, ces données sont depuis 2013 déclarées à la commune de chaque acheteur. Les dernières données disponibles au moment du projet « ARPEGES EDL 2019 » datent de 2015.

Depuis 2013, l'INRA travaille sur le développement d'une méthode de spatialisation de cette BNVD déclarée au code postal de l'acheteur (Marie Carles et al, 2015). Elle consiste à désagréger la quantité de produit, déclarée par code postal, entre tous les îlots culturels du Registre Parcellaire Graphique (RPG) et les îlots dits « OdR » (produits par l'INRA OdR de Toulouse pour compléter les parcelles et îlots concernés par des cultures non ou peu référencés par le RPG) associés à un même code postal.

La 1^{ère} étape est de calculer des coefficients de répartition du produit entre tous les îlots d'un code postal, en fonction des surfaces de chaque îlot et du groupe culturel de chaque îlot (pondération par la dose de référence du groupe culturel). La somme de tous les coefficients des îlots culturels rattachés à un même code postal doit faire 1. L'INRA (OdR de Toulouse) est en charge de ce travail. Il faut noter deux choses sur cette étape :

- Par cette méthode, on fait l'hypothèse que tous les îlots culturels rattachés à un même code postal sont traités selon le même Indice de Fréquence de Traitement (IFT) : la quantité appliquée sur chaque culture n'est pas pondérée par un usage éventuellement plus intensif (fréquence des applications selon les cultures) du produit sur telle ou telle culture.
- Les îlots culturels rattachés à un code postal ne sont pas forcément géographiquement contenus dans le code postal : la méthode tient compte du territoire d'exploitation, également connu via le RPG.

La 2^{ème} étape est de multiplier chaque coefficient d'îlot cultural par la quantité de produit déclarée pour le code postal auquel il est rattaché, puis de convertir les quantités de produit par îlot en quantités de substance active. L'Ineris est en charge de ce travail.

Bien que la majorité des substances ne soit pas appliquée sur l'ensemble des îlots culturaux français, le fait de travailler à l'échelle nationale rend ce travail de calcul très fastidieux. C'est pourquoi seule l'année 2015 de la BNVD a été utilisée pour produire les résultats du projet « ARPEGES EDL 2019 ».

Il serait intéressant d'appliquer le modèle sur plusieurs années, quand les données de la BNV-d spatialisée seront disponibles.

Si l'on travaille pour l'instant sur les données d'une année, une des limites de cette donnée est l'incertitude sur l'enchaînement « année d'achat – durée du stockage – année d'application ». En l'absence d'information, on fait l'hypothèse que l'année d'achat correspond à l'année d'application. Cependant, au moment du projet, les îlots OdR 2015 n'étaient pas encore disponibles. C'est pourquoi l'Ineris a réalisé le croisement des coefficients d'application 2015 sur le RPG complété des îlots OdR de l'année 2014.

Quantités orphelines

La BNVD spatialisée fournie par l'Ineris est accompagnée d'un bilan des quantités non spatialisées, appelées « quantités orphelines ». Le Tableau 5 présente ce bilan pour la France entière et pour les 15 substances actives étudiées dans le projet, en exprimant ce bilan en quantité et en ratio (quantités orphelines / quantités totales déclarées). On constate que trois substances sont spatialisées de façon très incomplète : l'imidaclopride (73,6% des quantités totales achetées n'ont pas pu être spatialisées), le métaldéhyde (54,6%) et l'oxadiazon (57%).

Substance active	Quantités orphelines (kg)	Ratio de quantités orphelines
2,4-d	58700	12,4%
2,4-mcpa	21782	3,4%
Aminotriazole	27623	7,4%
Bentazone	0	0,0%
Boscalid	16360	5,9%
Chlortoluron	62647	4,9%
Diflufénicanil	14289	3,5%
Glyphosate	393068	5,0%
Imidaclopride	191528	73,6%
Métaldéhyde	388007	54,6%
Métazachlore	195679	27,7%

Nicosulfuron	930	1,8%
Oxadiazon	1424	57,0%
Pendiméthaline	29853	2,8%
S-metolachlore	79718	4,0%

Tableau 5 : Quantités orphelines et ratios pour les substances du projet

L'INRA avance trois explications pour ces quantités orphelines :

- Le code postal acheteur n'a pas été renseigné : ce point n'est pas corrigible.
- L'Autorisation de Mise sur le Marché (AMM) manque dans les bases de données dont dispose l'INRA : la Dose Homologuée du produit, indiquée dans cet AMM, n'est donc pas disponible, ce qui empêche de spatialiser la quantité de produit achetée en question. Ce point peut être corrigé si l'AMM est trouvée.
- La non-spatialisation peut aussi s'expliquer par un décalage temporel (achats 2015 spatialisés sur l'occupation du sol 2014) ou spatial des données (approximations dans la localisation du siège d'exploitation par rapport à son parcellaire).

Ilots agricoles et non-agricoles de la BNVD spatialisée

La spatialisation de la BNVD concerne autant les îlots agricoles que les îlots non-agricoles. En effet, les îlots OdR complétant le RPG peuvent être agricoles (ex : vigne, qui apparaît très peu dans le RPG), ou non-agricoles (cimetières, terrains de sport, aérodromes, etc).

Pour ARPEGES, il faut donc soustraire les îlots non-agricoles, qui ne rentrent pas dans le périmètre du modèle. Les quantités appliquées sur îlots non-agricoles estimées par la spatialisation de la BNVD peuvent toutefois représenter une part importante des quantités spatialisées (Tableau 6), voire la totalité (cas du métaldéhyde).

Le fait que le métaldéhyde soit spatialisé uniquement sur îlots non-agricoles s'explique par le fait que tous les produits le contenant ont comme cible « limaces-escargots », qui est toujours associée à la culture « traitements généraux » dans la table des AMM. Or la culture « traitements généraux » est considérée comme relevant uniquement du non-agricole dans la méthode de spatialisation, exceptée pour le glyphosate.

La dernière colonne « quantités perdues lors de l'agrégation » vise à vérifier la conservation des quantités lors de l'agrégation. On relève une perte d'information très importante pour l'oxadiazon (58%) : les effets de bords de l'intersection entre les couches d'îlots agricoles et BVME sont suspectés, mais sont à confirmer.

Substance active	Nombre d'îlots agricoles	Quantités appliquées sur îlots agricoles (kg)	Nombre d'îlots non agricoles	Quantités appliquées sur îlots non agricoles (kg)	% quantités îlots agri/tot	Nombre de BVME faisant l'objet d'applications agricoles	Quantités appliquées sur les BVME faisant l'objet d'applications agricoles (kg)	Quantités perdues lors de l'agrégation
2,4-D	9 565 325	368 543	280 894	45 789	89%	9 543	363 911	1%
2,4-MCPA	8 443 066	61 7749	93 349	5 888	99%	9 409	611 219	1%
Aminotriazole	2 174 430	306 824	218 682	80 868	79%	5 944	302 116	2%
Bentazone	4 959 711	198 510	0	0	100%	7 964	191 337	4%
Bifenox	1 681 231	50 969	0	0	100%	5 927	49 276	3%
Boscalid	4 930 971	260 015	0	0	100%	8 510	256 527	1%
Chlortoluron	2 049 332	1 207 546	0	0	100%	7 584	1 195 806	1%
Diflufenicanil	2 580 944	384 124	268 302	10 189	97%	7 900	382 375	0%
Glyphosate	10 788 480	6 641 086	312 071	777 727	90%	9 547	6 555 256	1%
Imidaclopride	1 486 211	68 695	0	0	100%	5 068	68 259	1%
Metaldehyde	0	0	255 597	323 373	0%	0	0	
Metazachlore	2 370 301	509 463	0	0	100%	6 127	505 843	1%
Nicosulfuron	1 199 808	50 217	0	0	100%	7 717	49 775	1%
Oxadiazon	162 825	770	21 893	318	71%	1 196	326	58%
Pendimethaline	7 697 254	1 024 717	16 078	1 213	100%	8 784	1 017 507	1%
S-Métolachlore	3 094 555	1 929 853	0	0	100%	8 192	1 915 345	1%

Tableau 6 : Bilans sur les quantités agricoles, non-agricoles et les pertes de quantités après agrégation à l'échelle BVME

⇒ Coefficients de saisonnalité de l'application des substances actives :

Pour obtenir des pressions d'usage phytosanitaire en nappe haute et en nappe basse à partir des quantités annuelles de substance active, il faut appliquer des coefficients de saisonnalité à chaque substance.

Ces coefficients pourraient dans l'idéal s'appliquer à chaque îlot selon le groupe cultural pour les cas où il y aurait des différences de coefficient entre groupes culturaux, mais cela entraînerait un traitement supplémentaire beaucoup trop lourd. On applique donc dans Arpèges des coefficients « toutes cultures confondues » aux quantités agrégées par BVME, après l'étape d'ajustement à la

maille de calcul (voir 3.4.1.3). Les coefficients utilisés sont présentés ci-dessous mais à considérer comme une donnée de base à renseigner par expertise.

Ces coefficients peuvent être établis par expertise. A défaut, une méthode a été proposée (voir Annexe 1), des coefficients ont été établis par Irstea pour toute la France et ajustés pour le cas du bassin Seine-Normandie compte tenu d'une expertise locale qui a permis de moduler les coefficients proposés à l'échelle nationale. Ils sont présentés dans le Tableau 7 ci-dessous.

Substance active	Coefficients par défaut		Coefficients différents pour l'AESN	
	% en « nappe haute »	% en « nappe basse »	% en « nappe haute »	% en « nappe basse »
2,4-D	50	50		
2,4-MCPA	0	100		
Aminotriazole	50	50		
Bentazone	0	100		
Boscalid	0	100		
Chlortoluron	100	0		
Diflufenicanil	80	20	90	10
Glyphosate	50	50		
Métazachlore	60	40	90	10
Nicosulfuron	0	100		
Pendiméthaline	55	45		
S-métolachlore	0	100		

Tableau 7 : Coefficients de saisonnalité retenus pour le "projet ARPEGES EDL 2019"

Ce qui a changé depuis la méthode « ARPEGES – EDL 2013 » : Données d'entrée

- Données pour la pression d'usage phytosanitaire :

Dans la méthode « ARPEGES – EDL 2013 », la donnée d'entrée pour la pression d'usage phytosanitaire était le RICA (Réseau d'Information Comptable Agricole). Il s'agissait de la dépense en produits phytosanitaires (tous produits confondus) des exploitations, obtenue à l'échelle des Petites Régions Agricoles et datant de 2006.

Cette donnée était bien adaptée pour l'évaluation du potentiel de contamination par un cocktail de substances actives, mais ne suffisait plus pour une évaluation substance par substance. Il n'y avait par ailleurs qu'un seul couple de coefficients de saisonnalité, fixé à 50-50 % pour le cocktail de substance active : ces coefficients ont aussi dû être détaillés substance par substance.

Problématique

Le modèle ARPEGES demande en entrée 15 variables au total, dont 3 uniformes à l'échelle nationale (saison, Koc et DT50), et 12 qu'il faut renseigner à l'échelle du bassin versant de masse d'eau, unité spatiale du modèle. La France métropolitaine comporte environ 10 000 masses d'eau et les BVME ont en moyenne une superficie de 5 320 ha (4 fois la superficie moyenne d'une commune).

L'échelle spatiale des données disponibles est très variable. Par rapport à l'unité spatiale du modèle trois cas de figure se présentent (Figure 19) :

- 1^{er} cas : les entités des données peuvent être comprises dans les bassins versants locaux (exemples : mailles de l'IDPR, îlots culturels de la BNV-d spatialisée),
- 2^{ème} cas : les entités et les bassins versants peuvent se recouper (cantons ou unités cartographiques de sol par exemple)
- 3^{ème} cas : les entités peuvent être largement plus grandes que les bassins (zones climatiques par exemple).

Nous avons donc fait appel aux méthodes couramment utilisées pour aborder ces changements d'échelle (Vernier, Zahm et al. 2009) et ainsi permettre le croisement de l'ensemble des informations aux échelles adéquates pour ces trois cas de figure.

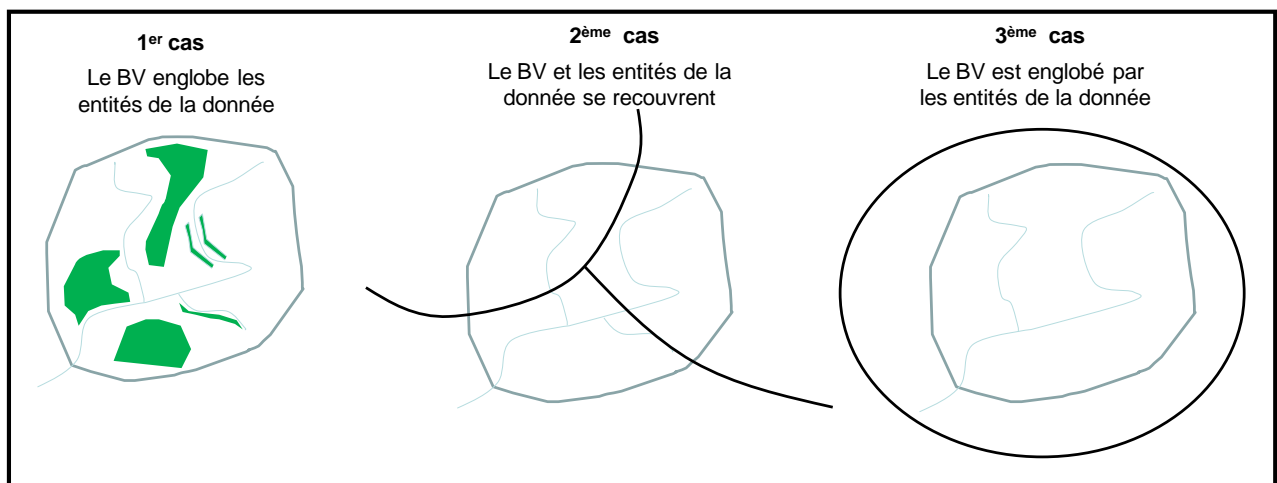


Figure 19 : Illustration des trois cas d'agrégation possibles

Le cas n°1 est lié aux données à plus grande échelle, le cas n°2 aux données à échelle équivalente et le cas n°3 aux données à plus petite échelle.

Méthodes d'agrégation*Cas n°1 : le BVME englobe les entités de la donnée*

Dans ce cas, l'entité spatiale s'organise selon une relation d'emboîtement⁸ avec l'entité spatiale BVME. Le BVME est le niveau supérieur englobant. Le transfert d'information s'opère donc par agrégation de l'information du niveau inférieur.

⁸ Dans une relation d'emboîtement le transfert d'informations s'opère le plus souvent au sein d'un même type d'organisation, entre entités spatiales emboîtées : de la parcelle à l'exploitation, de la commune au pays, ou encore du bassin versant au bassin versant englobant (Marceau, 1999 ; Dumanski et al., 1998)

Suivant la variable étudiée et la donnée utilisée pour estimer cette variable, on choisira la somme ou la moyenne comme opérateur mathématique pour l'agrégation.

Par exemple, pour estimer la variable « ripisylves 10 m » au niveau du BVME, on a choisi l'opérateur d'agrégation « somme », puisqu'on calcule la somme de tous les polygones BDTOPO végétation découpée dans la bande de 10 m de part et du réseau hydrographique parcourant le bassin versant d'une masse d'eau avant de la rapporter à la surface du bassin (cf. Annexe 8 pour plus de détail).

C'est aussi cette méthode qui est appliquée pour sommer les quantités annuelles de substance active par îlot de la BVND spatialisée.

Cas n°2 : le BVME et les entités de la donnée se recouvrent

Dans ce cas, l'entité spatiale de la donnée ne s'organise pas selon une relation d'emboîtement avec l'entité spatiale BVME car ces deux entités appartiennent à des niveaux d'organisation différents. Nous sommes donc confrontés à des entités spatiales en situation de recouvrement, c'est-à-dire qu'une partie seulement de l'entité spatiale est incluse dans une autre entité. Un BVME peut en effet se trouver sur le territoire d'une à plusieurs entités de la donnée, et à l'inverse, une entité de la donnée peut se partager entre plusieurs BVME.

Deux méthodes de transfert d'information entre deux espaces en situation de recouvrement ont été utilisées :

⇒ La méthode de désagrégation-réagrégation

Le transfert d'information entre deux espaces en situation de recouvrement impose généralement de se ramener à un *niveau d'échelle commun*, soit par agrégation à un niveau supérieur englobant les deux espaces, soit par désagrégation à un niveau inclus dans les deux espaces (Sanders, 2001). Pour la deuxième solution, l'étape initiale de désagrégation des données doit être suivie d'une démarche agrégative permettant de reconstituer l'information au niveau le plus pertinent.

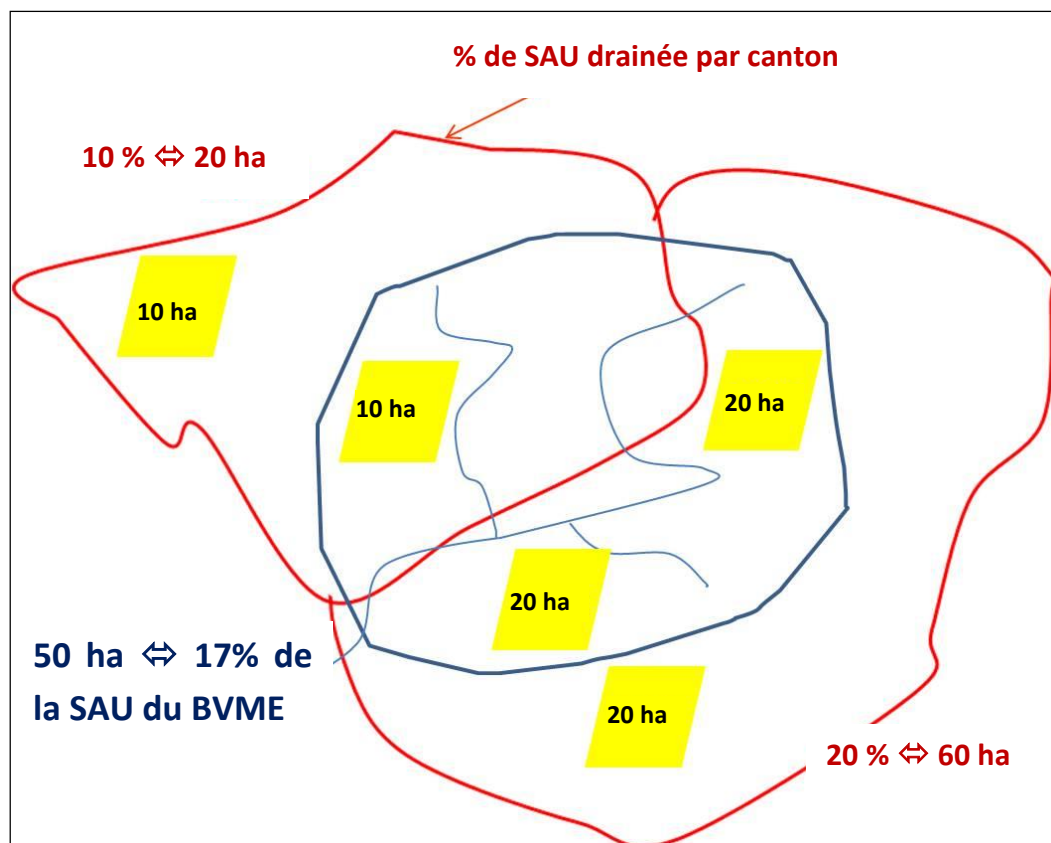
Dans notre cas, il n'existe pas de niveau d'échelle supérieur englobant et l'objectif de ce travail est de reconstituer l'information à un des deux niveaux d'échelle d'entrée : le BVME. On a donc opté pour la deuxième solution, la méthode de désagrégation-réagrégation, qui comporte deux étapes : (1) étape de désagrégation ; (2) étape de réagrégation.

Dans cette méthode, l'étape de désagrégation est cruciale. Elle nécessite généralement une donnée, appelée variable de contrôle, dont les entités sont incluses dans les deux espaces. Cette variable a pour but de répartir l'information de la donnée du niveau supérieur de la manière la plus pertinente possible pour l'agréger ensuite au niveau de l'entité BVME.

Ce type de méthode a été utilisé par exemple pour le transfert d'information des surfaces de drainage agricole à l'échelle des cantons. Ces surfaces ont été réparties sur les superficies agricoles (variable de contrôle) incluses dans une entité Canton et réaffectées au niveau des BVME en sommant les surfaces drainées des superficies agricoles incluses dans l'entité BVME (cf. Figure 20 et Annexe 4 pour plus de détail).

Cette méthode est utilisée lorsqu'on dispose d'une variable de contrôle pertinente pour l'étape de désagrégation sur l'ensemble du territoire métropolitain.

Figure 20 : Schématisation du calcul des surfaces de drainage agricole à l'échelle du BVME par la méthode de désagrégation-agrégation.



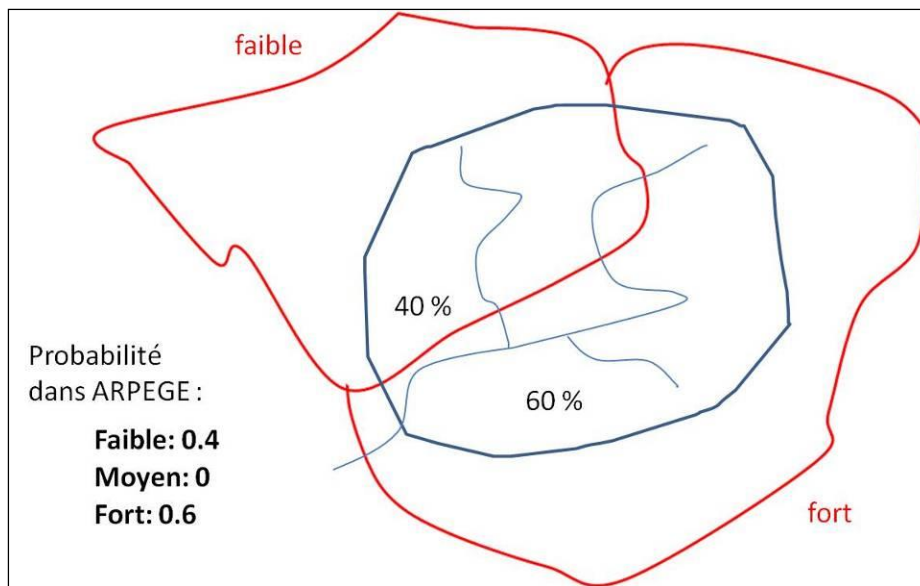
La variable de contrôle est la SAU (estimée à partir de CLC 2006, cf. Annexe 4). Les surfaces drainées à l'échelle des cantons sont tout d'abord ventilées en fonction de la SAU de chaque canton. Les surfaces drainées associées aux entités SAU intersectant l'entité BVME sont ensuite sommées pour calculer la SAU drainée à cette échelle. En rouge, les limites des cantons, en jaune les entités de SAU (variable de contrôle).

⇒ La méthode d'agrégation probabiliste

Cette méthode a été mise au point pour pallier l'absence d'une variable de contrôle pertinente (autre que la surface intersectée entre les BVME et la couche de données) dans la mise en œuvre de la méthode de désagrégation-réagrégation. Elle a été conçue en lien avec la technique de modélisation (réseau bayésien probabiliste, cf.3.3.2).

Dans cette méthode, la donnée pour estimer une variable est tout d'abord discrétisée, à son échelle d'origine, en fonction des classes d'état définies dans le modèle. Ensuite, les probabilités des classes d'état au sein d'un BVME correspondent au prorata des superficies des classes d'état de la donnée qui intersectent l'entité BVME (Figure 21).

Figure 21 : Schématisation de l'approche d'agrégation probabiliste.



Dans cet exemple la variable d'entrée (ex : la battance) dans le modèle a trois modalités : faible, moyen, fort. Les probabilités des classes d'état de la variable au sein d'une entité BVME sont fonction des superficies des classes d'état de la donnée qui intersectent cette entité. Ainsi, ce BVME aura une probabilité dans ARPEGES de 0,4 pour l'état « faible » car 40 % de sa superficie est recouverte par la donnée avec la classe « faible ». Même raisonnement, pour la probabilité de l'état moyen et de l'état fort. En rouge, les limites des entités de la donnée discrétisée en fonction des classes d'état définies dans le modèle.

Cas n°3 : le BVME est englobé par les entités de la donnée

Dans ce cas, l'entité spatiale s'organise selon une relation d'emboîtement avec l'entité spatiale BVME. L'entité de la donnée est le niveau supérieur englobant. Le transfert d'information s'opère donc par désagrégation de l'information du niveau supérieur.

La désagrégation de cette information est faite par jointure spatiale. Cette jointure permet de transférer l'information de la donnée du niveau supérieur à l'entité BVME selon une relation spatiale d'intersection.

Par exemple, pour estimer la variable « zone climatique » au sein de l'entité BVME, on utilise comme relation spatiale « ont leur centroïde dans ». L'entité BVME récupère alors comme valeur de zone climatique, la valeur de l'entité zone climatique qui contient son centroïde (cf. Annexe 10 pour plus de détail).

3.4.2. Intégration des données d'entrée : discrétisation des données d'entrée et traduction en n-uplets de probabilités

Une fois les données d'entrée collectées et agrégées à l'échelle de calcul d'ARPEGES, l'outil R discrétise lui-même les données continues selon les seuils de discrétisation évoqués au paragraphe 3.3.1.2. Il traduit ensuite la donnée discrétisée en un n-uplet de probabilités (n = nombre d'états pouvant être pris par la variable), qui est la donnée à proprement parler du réseau bayésien. Cela revient à affecter une probabilité de 100% à l'état que prend la donnée discrétisée sur une maille (ex : colonnes drainage et IDPR du Tableau 8). Si la donnée d'entrée est inconnue sur un BVME, l'outil

remplace la donnée manquante par un n-uplet équiprobable (faible : 33% ; moyen : 33, forte : 33%) sur le BVME correspondant. Le niveau de confiance associé est dans ce cas faible (voir 3.5.1.3)

Deux exceptions existent pour les données d'entrée continues : dans le cas des bandes enherbées et du cumul de pluie, la traduction en doublet correspond à une normalisation de la donnée d'entrée continue (voir colonne cumul_pluies du Tableau 8).

Dans le cas des données discrètes dès la collecte, dont les données de la BDGSF (voir colonne battance dans le Tableau 8), l'agrégation probabiliste a déjà produit des n-uplets de probabilités (cf 3.4.1.3).

Les trois options pour traduire les données d'entrée en n-uplets sont récapitulées donnée par donnée dans la dernière colonne du Tableau 3.

IDBV	Battance	Drainage	saison	Zone climato	cumul pluie	IDPR	DT50	KOC
BV100	{faible .42, moyen .58, fort 0}	{faible 1, moyen 0, fort 0}	nappe_basse	Mediterraneeen_sec	{faible 65.57, fort 34.43}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV1000	{faible .75, moyen 0, fort .25}	{faible 0, moyen 0, fort 1}	nappe_basse	Continental_Est	{faible 73.1, fort 26.9}	{faible 0, moyen 1, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10000	{faible 0, moyen .74, fort .26}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 78.47, fort 21.53}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10001	{faible .11, moyen .05, fort .84}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 65.83, fort 34.17}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10002	{faible .41, moyen .17, fort .42}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 66.31, fort 33.69}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10003	{faible .43, moyen .42, fort .15}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 58.04, fort 41.96}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10004	{faible .37, moyen 0, fort .63}	{faible 0, moyen 1, fort 0}	nappe_basse	Sud_Ouest	{faible 78.42, fort 21.58}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10005	{faible .93, moyen .07, fort 0}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 72.75, fort 27.25}	{faible 0, moyen 1, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10006	{faible .23, moyen .26, fort .51}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 70.13, fort 29.87}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10007	{faible .12, moyen .22, fort .67}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 57.93, fort 42.07}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10008	{faible .6, moyen 0, fort .4}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 71.09, fort 28.91}	{faible 0, moyen 1, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10009	{faible .55, moyen .1, fort .34}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 70.92, fort 29.08}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV1001	{faible 0, moyen .75, fort .25}	{faible 1, moyen 0, fort 0}	nappe_basse	Corridor_rhodanien	{faible 47.4, fort 52.6}	{faible 1, moyen 0, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10010	{faible .05, moyen .31, fort .64}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 73.69, fort 26.31}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10011	{faible .74, moyen .04, fort .22}	{faible 0, moyen 1, fort 0}	nappe_basse	Sud_Ouest	{faible 77.43, fort 22.57}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10012	{faible .08, moyen .47, fort .45}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 70.31, fort 29.69}	{faible 0, moyen 1, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10013	{faible .75, moyen .1, fort .16}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 77.39, fort 22.61}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10014	{faible .59, moyen .08, fort .33}	{faible 0, moyen 1, fort 0}	nappe_basse	Sud_Ouest	{faible 78.27, fort 21.73}	{faible 0, moyen 1, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10015	{faible .71, moyen 0, fort .29}	{faible 0, moyen 1, fort 0}	nappe_basse	Sud_Ouest	{faible 76.4, fort 23.6}	{faible 0, moyen 1, fort 0}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10016	{faible 0, moyen .68, fort .32}	{faible 1, moyen 0, fort 0}	nappe_basse	Sud_Ouest	{faible 75.29, fort 24.71}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}
BV10017	{faible .49, moyen .32, fort .19}	{faible 0, moyen 1, fort 0}	nappe_basse	Sud_Ouest	{faible 80.76, fort 19.24}	{faible 0, moyen 0, fort 1}	{courte 0, moyenne 1, longue 0}	{faible 0, moyen 1, fort 0}

Tableau 8 : Extrait de la table d'entrée pour le S-métolachlore en nappe basse

Les possibilités de l'interface R : Données d'entrée

- Import possible de nouvelles données d'entrée : attention, celles-ci peuvent nécessiter dans certains cas de revoir les seuils, ce qui est à faire avec précaution.
- Si un import de nouvelle donnée est réalisé, les étapes revenant à l'utilisateur sont la collecte, l'agrégation à l'échelle des BVME selon la méthode adéquate. La traduction en n-uplets de probabilités est réalisée par le programme associé à l'interface R.
- Des cartes d'indice de confiance et de probabilité de la modalité fort est également possible pour les données d'entrée (cf 3.5.1.3).

3.5. Types de sorties du réseau bayésien d'ARPEGES

Dans les sorties d'ARPEGES, on distingue d'abord les sorties brutes, correspondant à une table des résultats pour chaque BVME et chaque saison, puis les sorties retravaillées (extraction de grandeurs interprétables) et les cartes de résultats une fois la table pré-traitée devenue table attributaire de la couche SIG des BVME. Enfin, parmi ces résultats, on obtient des résultats communs à toutes les substances actives.

3.5.1. Sorties brutes

3.5.1.1. Fichiers de sortie par saison

Tout comme on envoie en entrée une table par saison (probabilité de 100% sur nappe haute ou nappe basse), on obtient en sortie une table par saison, résultant du calcul des probabilités des nœuds enfants à travers le réseau. Bien que la structure graphique du réseau présente 4 nœuds pour les potentiels de contamination lorsque la saison est équiprobable (cf Figure 15), il ne faut considérer que les potentiels de contamination de la saison considérée dans les résultats possibles (cf Figure 22 pour nappe basse, Figure 23 pour nappe haute).

Le résultat le plus brut en sortie de réseau est un n-uplet de probabilités pour chaque nœud enfant (n = nombre d'états du nœud enfant) et chaque BVME. Un extrait de la table pour le S-métolachlore en nappe basse est présenté dans le Tableau 9 (le tableau a été tronqué sur les lignes mais aussi sur les colonnes). Chaque colonne correspond à un nœud enfant, chaque ligne à un BVME.

IDnum	probabilities of vul_intr_ruissell	probabilities of vul_intr_subsurf	probabilities of vul_intr_drainage
101	(0 0.05688 0.94312)	(0.7 0.3 0)	(1 0 0)
1001	(0.49125 0.25875 0.25)	(1 0 0)	(0.35 0.5 0.15)
9995	(0.0027508 0.289227 0.708023)	(0.56 0.44 0)	(0.834 0.166 0)
9996	(0.047301 0.556311 0.396388)	(0.570297 0.429703 0)	(0.844554 0.155446 0)
9997	(0.06984 0.615774 0.314386)	(0.508 0.492 0)	(0.804 0.196 0)
9998	(0.018105 0.47422 0.507675)	(0.504 0.496 0)	(0.802 0.198 0)
9999	(0.0336583 0.482751 0.483591)	(1 0 0)	(0.329 0.597 0.074)
10000	(0.540088 0.439368 0.0205439)	(0.826 0.174 0)	(0.926 0.074 0)
10001	(0.07404 0.523337 0.402623)	(0.581 0.419 0)	(0.852 0.148 0)
10002	(0.0157327 0.452397 0.53187)	(0.514 0.486 0)	(0.807 0.193 0)
10003	(0.531 0.46 0.009)	(0.7 0.3 0)	(0.8 0.2 0)
10004	(0.0839638 0.649077 0.266959)	(0.557 0.443 0)	(0.838 0.162 0)
1002	(0.55 0.325 0.125)	(0.866 0.134 0)	(1 0 0)
10005	(0.006557 0.367288 0.626155)	(0.589 0.411 0)	(0.849 0.151 0)
10006	(0.032412 0.657028 0.31056)	(1 0 0)	(0.227 0.625 0.148)
10007	(0.3678 0.528369 0.103831)	(0.726 0.274 0)	(0.813 0.187 0)

Tableau 9 : Extrait d'une table de résultat (S-métolachlore nappe basse)

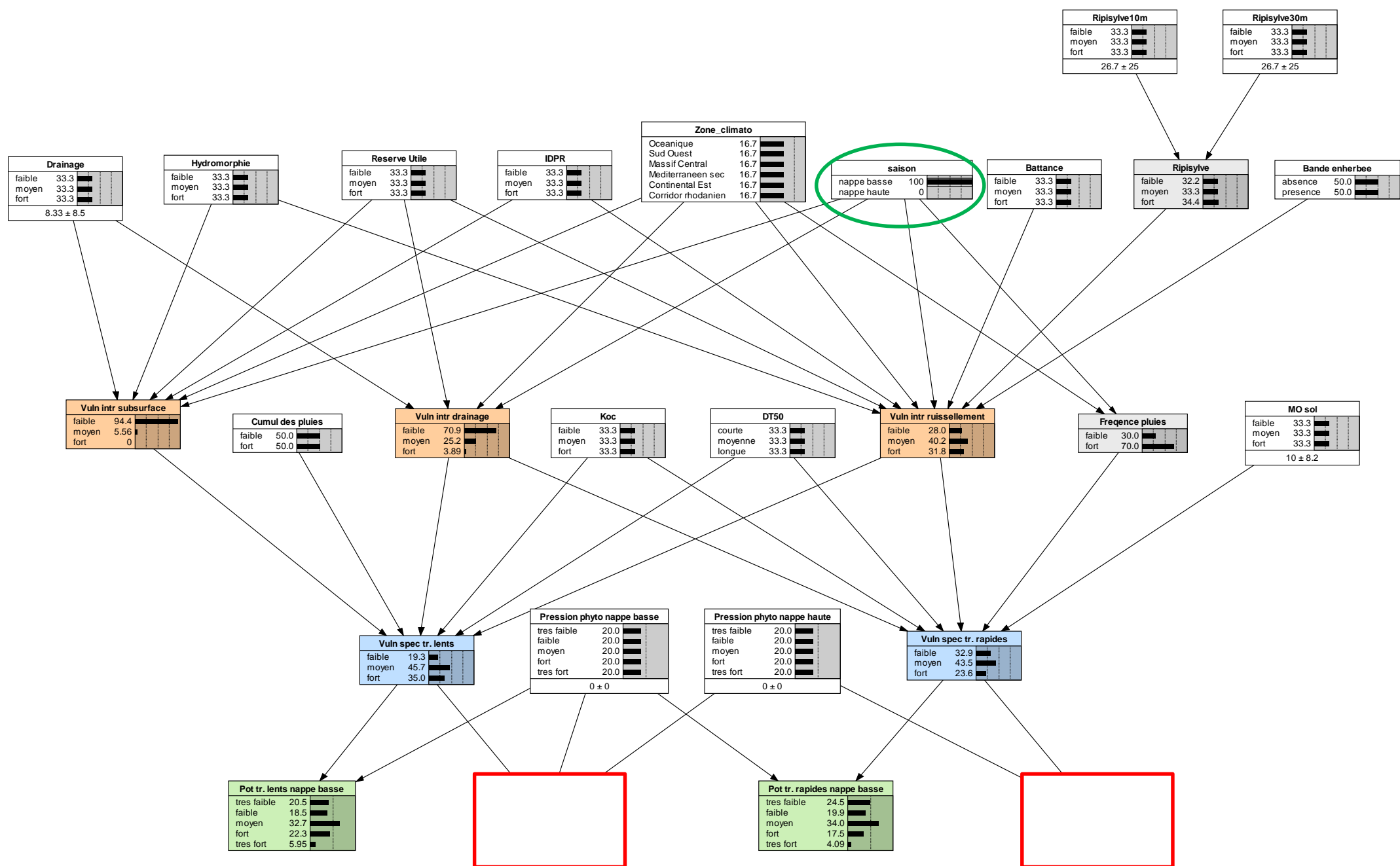


Figure 22 : Graphe du réseau ARPEGES en saison nappe basse

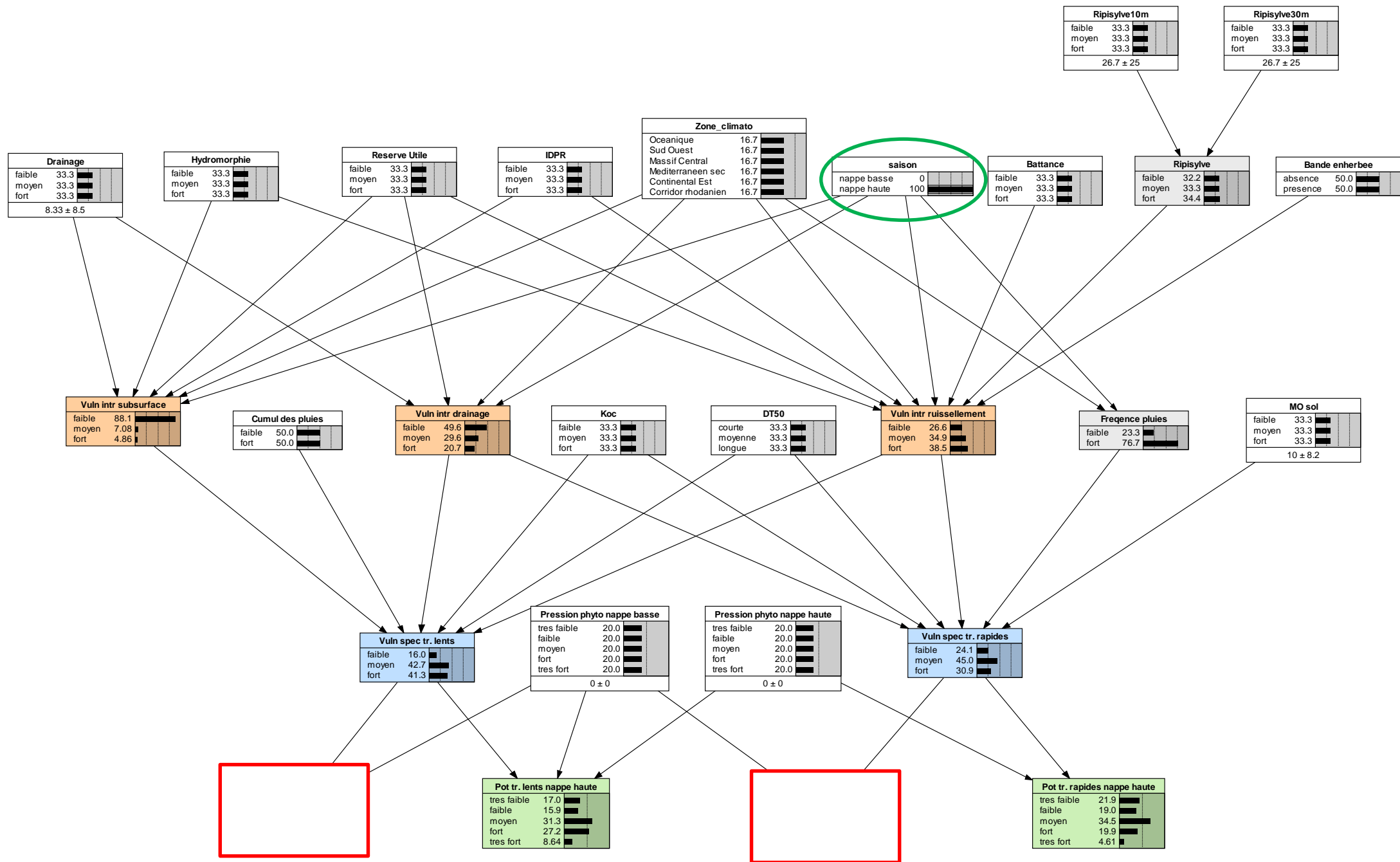


Figure 23 : Graphe du réseau ARPEGES en saison nappe haute

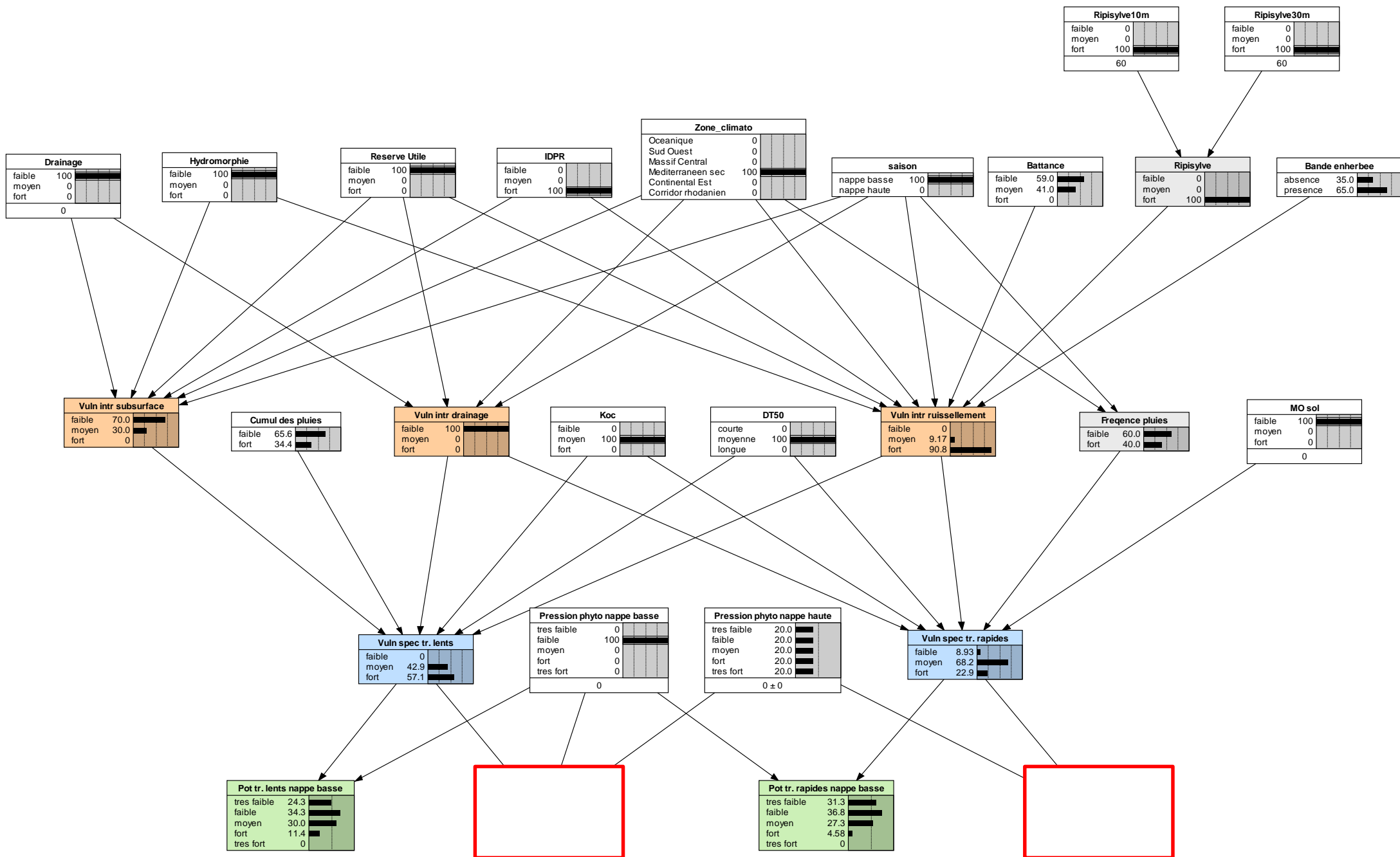


Figure 24 : Illustration sur le graphique du réseau ARPEGES du calcul pour un BVME exemple (valeurs des nœuds parents renseignées)

3.5.1.2. Lecture des n-uplets de probabilités et calcul opérés pour obtenir ces n-uplets

De façon générale, les nœuds enfants correspondant aux vulnérabilités intrinsèques et spécifiques représentent le niveau de vulnérabilité du bassin versant. Ce niveau de vulnérabilité peut être faible, moyen ou fort.

De même, les nœuds représentant les potentiels de contamination via les transferts lents ou rapides, codent ce niveau de risque en très faible, faible, moyen, fort ou très fort.

Les probabilités associées à chacun de ces niveaux de vulnérabilité ou de potentiel de contamination expriment le risque associé à ces niveaux. Par exemple, pour le S-métolachlore en nappe basse et le BVME d'identifiant IDnum = 101 (Tableau 9), si la vulnérabilité intrinsèque aux écoulements de subsurface a une probabilité de 70% d'être dans la classe « faible » (et de 30% d'être dans la classe moyenne), on parlera d'un risque élevé (équivalent à une forte probabilité) d'avoir une vulnérabilité faible aux écoulement de subsurface (autrement dit, on a de fortes chances que le bassin versant soit faiblement sensible aux écoulements de subsurface). Si c'était la classe « fort » qui avait une probabilité de 70%, on aurait un risque élevé (équivalent à une forte probabilité) d'avoir une vulnérabilité forte aux écoulements de subsurface.

Pour connaître le principe des calculs conduisant à ces n-uplets de probabilités, se référer à l'encadré « Les fondements mathématiques 3 ».

En pratique, on peut visualiser sur le graphique du réseau ARPEGES le résultat de la cascade de calcul pour un BVME exemple : reprenons celui dont l'IDnum = 101. Le Tableau 10 rappelle les données d'entrée pour cet exemple, le Tableau 11 ses données de sortie. Sur la Figure 24, on retrouve bien ces mêmes données d'entrée et de sortie, aux approximations près imposées par la précision des curseurs manuels dans l'outil Netica.

BV100	Entrées
Idnum 101	
Hydromorphie	{faible 1, moyen 0, fort 0}
RU	{faible 1, moyen 0, fort 0}
Battance	{faible .42, moyen .58, fort 0}
Drainage	{faible 1, moyen 0, fort 0}
saison	printemps_ete
Zone_climato	Mediterraneen_sec
cumul_pluie	{faible 65.57, fort 34.43}
IDPR	{faible 0, moyen 0, fort 1}
Ripisylve10m	{faible 0, moyen 0, fort 1}
Ripisylve30m	{faible 0, moyen 0, fort 1}
bande_enherbee	{presence .36, absence .64}
MO	{faible 1, moyen 0, fort 0}
DT50	{courte 0, moyenne 1, longue 0}
KOC	{faible 0, moyen 1, fort 0}
Pression_phyto_nappe_basse	{tres_faible 0, faible 1, moyen 0, fort 0, tres_fort 0}

Tableau 10 : Données d'entrée pour le BVME exemple (IDnum = 100, S-métolachlore nappe basse)

IDnum	Sorties
probabilities of vul_intr_ruissell	(0 0.06 0.94)
probabilities of vul_intr_subsurf	(0.7 0.3 0)
probabilities of vul_intr_drainage	(1 0 0)
probabilities of vul_spe_aigue	(0.08 0.69 0.23)
probabilities of vul_spe_chro	(0 0.42 0.58)
probabilities of risque_cont_aigue_ete	(0.31 0.37 0.28 0.05 0)
probabilities of risque_cont_chro_ete	(0.24 0.34 0.3 0.12 0)
probabilities of risque_cont_aigue_hiver	(0.18 0.18 0.40 0.20 0.04)
probabilities of risque_cont_chro_hiver	(0.13 0.14 0.32 0.32 0.09)

Tableau 11 : Données de sortie pour le BVME exemple (IDnum = 100, S-métolachlore nappe basse)

3.5.1.3. Interprétation des n-uplets : probabilité majoritaire et indice de confiance

Pour chaque nœud à prédire (nœuds enfant pour ARPEGES), on obtient une distribution de probabilités prédites, c'est-à-dire une probabilité pour chacun des états possibles d'un nœud. On peut donc déterminer l'état le plus probable qui est l'état pour le nœud concerné qui a la plus forte probabilité d'occurrence, dans le cas envisagé. Cette information est utile mais insuffisante : il s'agit d'une vision assez réductrice et résumée des résultats.

Il est intéressant de conserver le lien avec la distribution complète, car elle donne une information complémentaire très importante : le **niveau de confiance** associé au résultat.

En effet, que l'état le plus probable ait une probabilité d'occurrence très élevée ne signifie pas la même chose que si deux ou trois états ont des probabilités sensiblement identiques. Dans le premier cas, on accorde un niveau de confiance élevé au résultat, tandis que dans le second, il est plus difficile de conclure ou d'accepter la prédiction.

Afin d'estimer ce niveau de confiance, une façon simple de procéder est de considérer les probabilités associées aux deux classes les plus probables et de calculer l'écart entre ces probabilités. Si cet écart est important, alors la confiance dans la prédiction est plus forte car on hésite moins entre les deux classes possibles. Au contraire, si l'écart est faible, le risque de se tromper augmente (dans le cas extrême, on revient au tirage pile ou face, où chaque événement a une probabilité de 50%. On a alors une chance sur deux de se tromper en pariant sur l'un des événements).

On obtient alors une table de sortie retravaillée (ex : Tableau 9 retravaillé et présenté en

Tableau 12), avec pour chaque variable de sortie, une colonne pour la probabilité majoritaire dans le n-uplet, une colonne pour la valeur de probabilité de chaque état de la variable (séparation du n-uplet) et une colonne pour l'indice de confiance de cette variable.

IDnum	proba_maj_ruiss	vuln_ruiss_faible	vuln_ruiss_moyen	vuln_ruiss_fort	IC_ruiss	proba_maj_sub	vuln_sub_faible
101	fort	0	0.05688	0.94312	0.88624	faible	0.7
1001	faible	0.49125	0.25875	0.25	0.2325	faible	1
9995	fort	0.0027508	0.289227	0.708023	0.418796	faible	0.56
9996	moyen	0.047301	0.556311	0.396388	0.159923	faible	0.570297
9997	moyen	0.06984	0.615774	0.314386	0.301388	faible	0.508
9998	fort	0.018105	0.47422	0.507675	0.033455	faible	0.504
9999	fort	0.0336583	0.482751	0.483591	0.000840000000000007	faible	1
10000	faible	0.540088	0.439368	0.0205439	0.10072	faible	0.826
10001	moyen	0.07404	0.523337	0.402623	0.120714	faible	0.581
10002	fort	0.0157327	0.452397	0.53187	0.079473	faible	0.514
10003	faible	0.531	0.46	0.009	0.071	faible	0.7
10004	moyen	0.0839638	0.649077	0.266959	0.382118	faible	0.557
1002	faible	0.55	0.325	0.125	0.225	faible	0.866
10005	fort	0.006557	0.367288	0.626155	0.258867	faible	0.589
10006	moyen	0.032412	0.657028	0.31056	0.346468	faible	1
10007	moyen	0.3678	0.528369	0.103831	0.160569	faible	0.726
10008	moyen	0.0267327	0.675386	0.297881	0.377505	faible	0.556
10009	moyen	0.321128	0.384579	0.294294	0.063451	faible	1

Tableau 12 : Extrait du tableau 6 retravaillé

Les fondements mathématiques 3 : Entrées et sorties d'un réseau bayésien

Une fois le modèle écrit et les tables de probabilités conditionnelles remplies, il est possible d'utiliser un réseau de probabilités bayésien comme un modèle prédictif. Il s'agit alors de fixer l'état de certains nœuds afin de prédire l'état d'autres nœuds pour lesquels on n'a pas d'observation.

Il est intéressant de noter que cette prédiction peut, de manière générale, aller aussi bien dans un sens descendant (les causes du phénomène sont fixées, le modèle prédit les conséquences) que dans un sens ascendant (les conséquences sont observées et donc fixées, le modèle prédit les causes de ces observations).

Cependant, l'utilisation qui en est faite dans la méthode ARPEGES est la première option uniquement : on utilise le réseau bayésien en fixant les probabilités d'états des nœuds parents « d'entrée » pour obtenir les probabilités d'états des nœuds enfants.

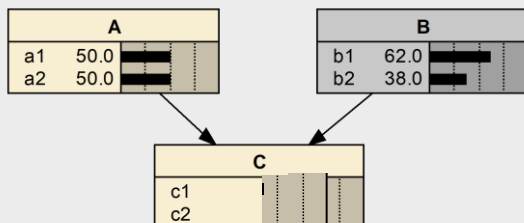
- Probabilités d'états des nœuds parents

On distingue les nœuds parents « intermédiaires », qui sont aussi des nœuds enfants issus d'un croisement de variables dans le réseau bayésien, des nœuds parents « d'entrée » pour lesquels il faut fournir des données.

Pour les nœuds parents « d'entrée », deux cas de figure se présentent :

- ⇒ Soit on dispose de l'information pour la variable, de manière plus ou moins certaine : dans ce cas, on traduit cette information en distribution de probabilités pour les états de la variable.
- ⇒ Soit on ne dispose d'aucune information sur la variable : dans ce cas, le réseau bayésien traduit automatiquement ceci en distribution uniforme de probabilité entre les états du nœud inconnu.

Reprenons l'exemple simple présenté dans l'encadré précédent d'un modèle à trois nœuds, dont deux parents et un enfant, chacun ayant deux états possibles : A (a1, a2), B (b1, b2) et C (c1, c2).



Si on ne connaît pas l'état de A, comme il y a deux états pour cette variable, la distribution de probabilités est de 50 % pour chaque état a1 et a2, soit (0.5 , 0.5).

Si, pour la variable B, la probabilité d'occurrence de b1 est de 62%, par conséquent celle de b2 est de 38% et la distribution de probabilités est (0.62 , 0.38). On peut aussi représenter graphiquement ces distributions avec le logiciel NETICA, ce qui donne la figure ci-contre.

Le paragraphe suivant explique comment on obtient la distribution de probabilités pour les états (c1 , c2).

Pour les nœuds parents « intermédiaires », les distributions de probabilité de leurs états sont calculées comme des nœuds enfants selon les propriétés du réseau bayésien (explication ci-après).

- Calcul prédictif de probabilités d'états des nœuds enfants

Les calculs prédictifs reposent intégralement sur les propriétés des probabilités conditionnelles. Le résultat de ces calculs est en fait le barycentre de l'ensemble des cas possibles, c'est-à-dire la somme des probabilités des états du nœud enfant conditionnellement à l'ensemble des combinaisons d'états de ses nœuds parents pondérées par les probabilités respectives d'occurrence des combinaisons.

Toujours sur l'exemple précédent : si l'on cherche à prédire la distribution de probabilités des états du nœud C, on peut décomposer la probabilité que C prenne l'état c1 selon les différentes combinaisons (4 au total) des états de A et B :

$$\begin{aligned} P(C=c1) = & p(C=c1 \cap A=a1 \cap B=b1) + \\ & p(C=c1 \cap A=a1 \cap B=b2) + \\ & p(C=c1 \cap A=a2 \cap B=b1) + \\ & p(C=c1 \cap A=a2 \cap B=b2) \end{aligned}$$

Les fondements mathématiques : Entrées et sorties d'un réseau bayésien (suite)

Les propriétés des probabilités conditionnelles permettent d'écrire ce type d'égalité :

$$p(C=c1 \cap A=a1 \cap B=b1) = p(C=c1 | A=a1 \cap B=b1) \cdot p(A=a1 \cap B=b1)$$

Ainsi, on obtient :

$$\begin{aligned} P(C=c1) &= p(C=c1 | A=a1 \cap B=b1) \cdot p(A=a1 \cap B=b1) + \\ &\quad p(C=c1 | A=a1 \cap B=b2) \cdot p(A=a1 \cap B=b2) + \\ &\quad p(C=c1 | A=a2 \cap B=b1) \cdot p(A=a2 \cap B=b1) + \\ &\quad p(C=c1 | A=a2 \cap B=b2) \cdot p(A=a2 \cap B=b2) \end{aligned}$$

Les variables A et B étant indépendantes, on peut ensuite écrire cette formule, dont on connaît tous les termes :

$$\begin{aligned} P(C=c1) &= p(C=c1 | A=a1 \cap B=b1) \cdot p(A=a1) \cdot p(B=b1) + \\ &\quad p(C=c1 | A=a1 \cap B=b2) \cdot p(A=a1) \cdot p(B=b2) + \\ &\quad p(C=c1 | A=a2 \cap B=b1) \cdot p(A=a2) \cdot p(B=b1) + \\ &\quad p(C=c1 | A=a2 \cap B=b2) \cdot p(A=a2) \cdot p(B=b2) \end{aligned}$$

A	B	c1	c2
a1	b1	40	60
a1	b2	80	20
a2	b1	50	50
a2	b2	70	30

On a la formule analogue pour $P(C=c2)$.

- Si l'on reprend les distributions de probabilité données en exemple précédemment pour A et B (où A n'est pas connu et B a plus de probabilité d'être en état b1), on obtient le processus de calcul et la distribution de probabilités pour C suivants:

A	p(A=...)	B	p(B=...)
a1	0,5	b1	0,62
a1	0,5	b2	0,38
a2	0,5	b1	0,62
a2	0,5	b2	0,38

	p(C=c1 A=... ∩ B=...)	p(C=c2 A=... ∩ B=...)
a1 b1	0,4	0,6
a1 b2	0,8	0,2
a2 b1	0,5	0,5
a2 b2	0,7	0,3

Calculs pour chaque combinaison

p(C=c1 A=... ∩ B=...) . p(A=...) . p(B=...)	p(C=c2 A=... ∩ B=...) . p(A=...) . p(B=...)
0,124	0,186
0,152	0,038
0,155	0,155
0,133	0,057

p(C=c1) = somme colonne = 0,564
p(C=c2) = somme colonne = 0,436

C	
c1	56,4
c2	43,6

- Si on est certain que l'état de A est a1 et que l'état de B est b2, alors leurs distributions de probabilités respectives sont (1,0) et (0,1), et on obtient comme résultat pour C les probabilités que l'on pouvait lire dans la table de probabilité conditionnelle de C pour le cas (A=a1, B=b2) :

$$P(C=c1) = p(C=c1 | A=a1, B=b2) = 80 \%$$

$$P(C=c2) = p(C=c2 | A=a1, B=b2) = 20 \%$$

3.5.2. Présentation des résultats sous forme de cartes

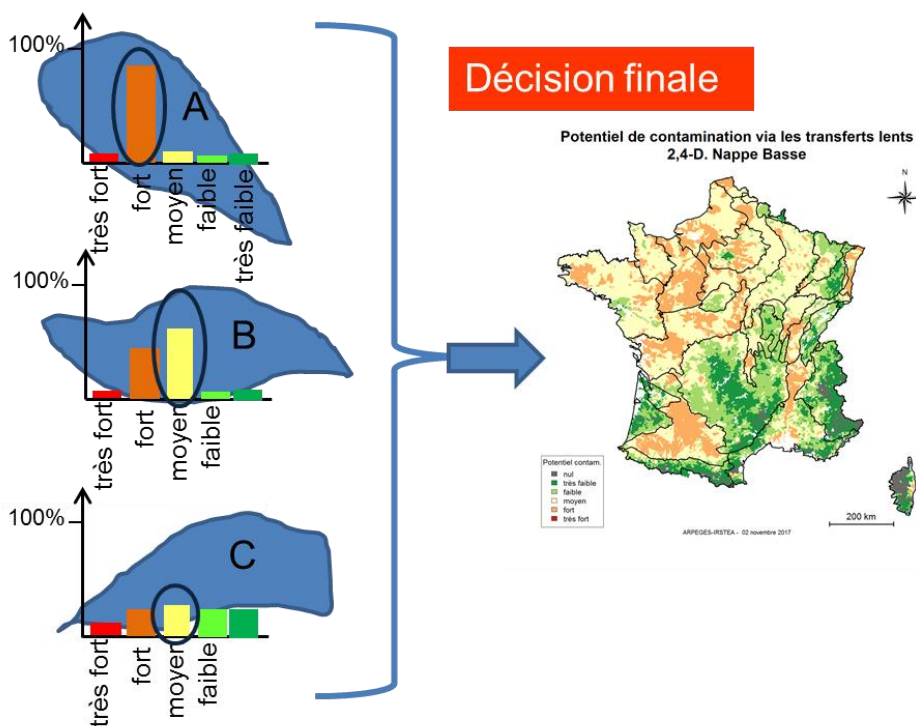
Les sorties retravaillées constituent une table attributaire de la couche d'information géographique des BVME. On peut ainsi faire une analyse thématique sur une colonne à la fois, chaque ligne correspondant à un polygone de BVME. Les types de cartes pouvant être produites sont donc, pour chaque variable de sortie :

⇒ Cartes de décision finale (= probabilité majoritaire)

Une colonne de sortie est la décision finale. Cela correspond, pour chaque ligne de la table et chaque polygone de la carte, à la classe de la probabilité la plus élevée du n-uplet de la variable cartographiée (cercles noirs sur le schéma ci-dessous).

Exemple de 3 BVME fictifs

→ Carte obtenue



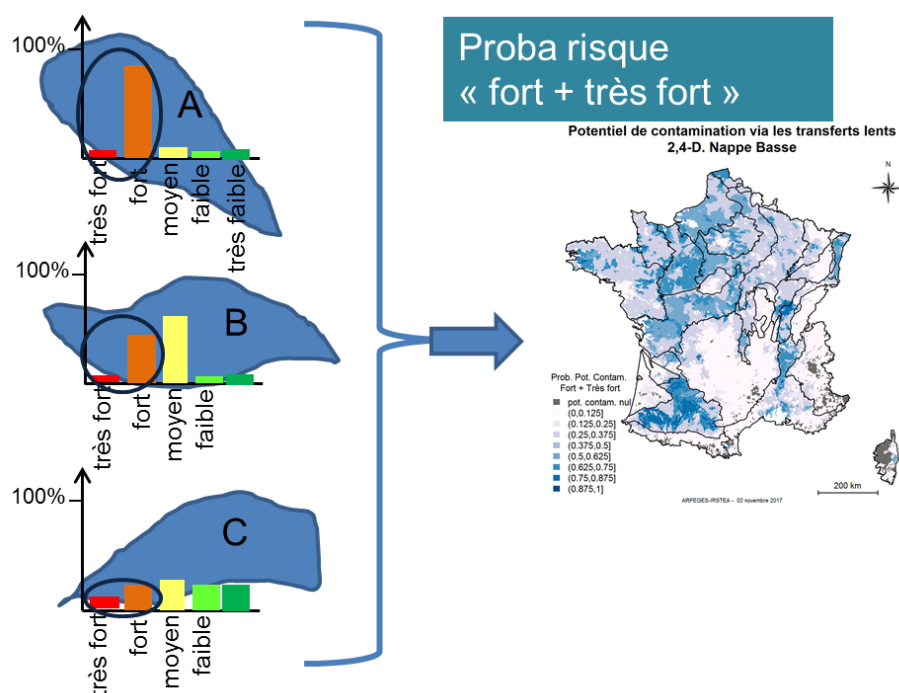
⇒ Cartes de probabilité d'un ou plusieurs états

La table de sortie contient pour chaque variable autant de colonne que le nombre de classes de cette variable. Cela revient, pour chaque ligne de la table et chaque polygone de la carte, à la valeur de la probabilité (cercles noirs sur le schéma ci-dessous) de la classe de la variable que l'on souhaite cartographier (ou de la somme des probabilités de plusieurs classes).

Exemple de 3 BVME fictifs



Carte obtenue



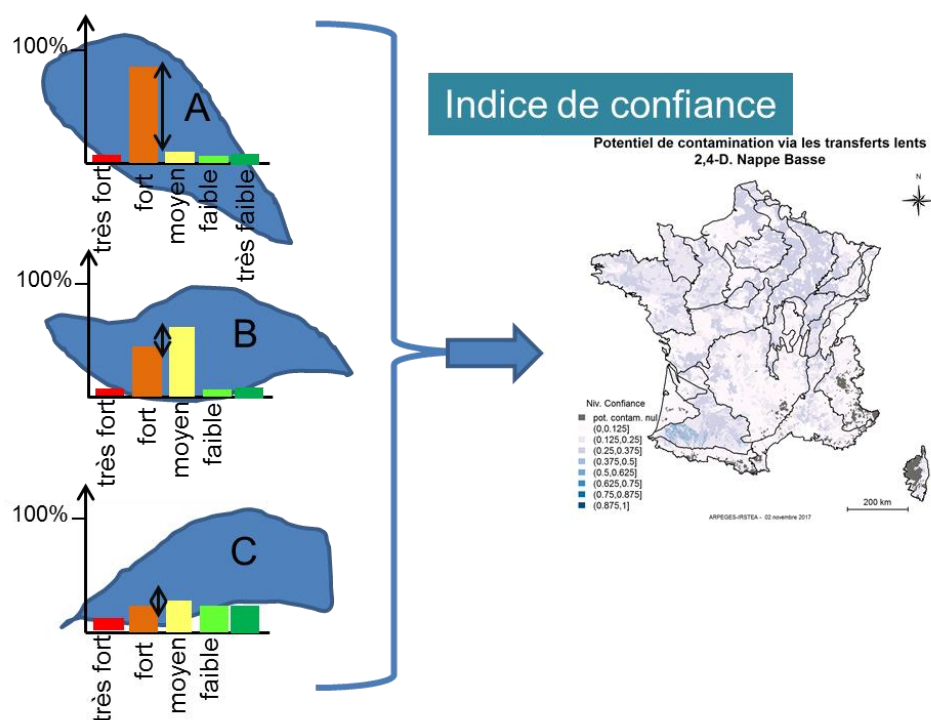
⇒ **Cartes d'indice de confiance**

Cela consiste, pour chaque ligne de la table de sortie et chaque polygone de la carte, à afficher l'indice de confiance associé à la variable que l'on souhaite cartographier (doubles-flèches noires sur le schéma ci-dessous).

Exemple de 3 BVME fictifs



Carte obtenue



L'interprétation des résultats d'ARPEGES consiste en le recouplement de ces trois types de sorties. Sur l'exemple des 3 BVME fictifs des schémas ci-dessus, on peut par exemple dire que :

- le BVME « A » a un potentiel de contamination « fort » avec un indice de confiance élevé ;
- le BVME « B » a un potentiel de contamination « moyen » avec un faible indice de confiance, et un potentiel de contamination « tangent » entre les classes « moyen » et « fort » (probabilités proches de 50% pour chacune de ces classes) ;
- le BVME « C » a un potentiel de contamination « moyen » également avec un faible indice de confiance, mais avec une quasi-équiprobabilité entre chaque classe de potentiel.

BVME fictif	Décision finale	Probabilité fort+très fort	Indice de confiance
A	Fort	Env. 80%	Env. 70%
B	Moyen	Env. 50%	Env. 10%
C	Moyen	Env. 20%	Env. 5%

Les possibilités de l'interface R : Cartographie des résultats

Possibilité de cartographier à d'autres échelles (Agence, groupe de BVME,...)

Export en shp ou csv pouvant faire l'objet d'une jointure dans un SIG – idem pour données d'entrée

3.5.3. Sorties communes à toutes les substances actives

Il est intéressant de noter que certaines sorties, et par conséquent certaines cartes, sont identiques d'une substance active à l'autre.

Cartes de vulnérabilités intrinsèques (

Figure 25,

⇒ Figure 26 et Figure 27 ; à gauche la carte de probabilité majoritaire, à droite l'indice de confiance associé).

⇒ Cartes de vulnérabilités spécifiques :

Etant donné que chaque caractéristique de substance active est discrétisée en 3 classes, il y a 9 combinaisons possibles de ces classes, pouvant être associées à des « substances –types », et par conséquent, 9 cartes possibles pour chaque vulnérabilité spécifique. Un exemple pour la vulnérabilité spécifique via les transferts lents en nappe basse est présenté ci-après, les autres sont à trouver en Annexe 14.

Ces cartes permettent déjà de dégager spatialement des comportements par famille de substances actives. Le

Tableau 13 reporte les quinze substances actives étudiées dans le projet « ARPEGES EDL 2019 » dans ce modèle de compilation des 9 cartes de chaque vulnérabilité spécifique.

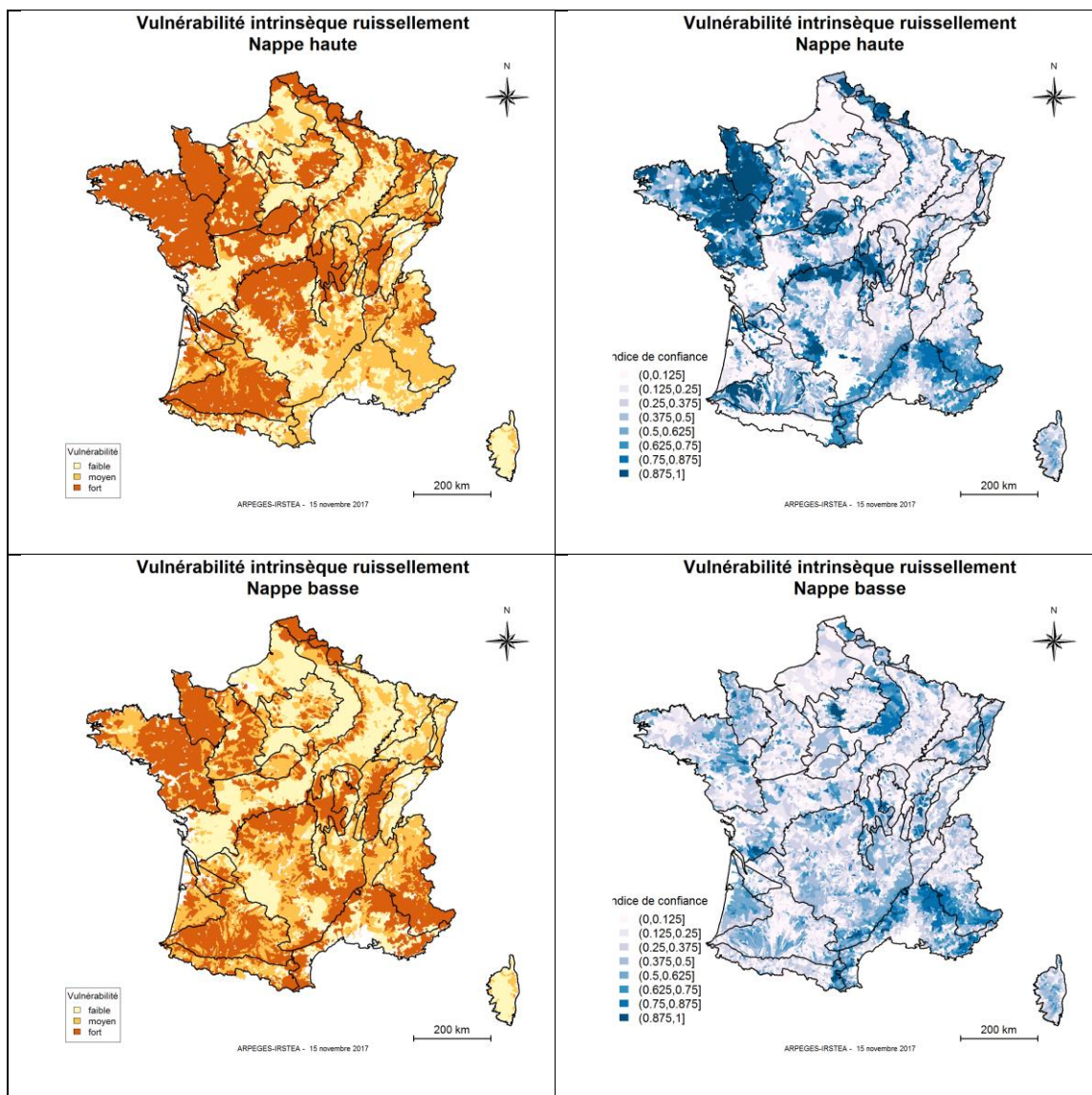


Figure 25 : Cartes de vulnérabilité intrinsèque au ruissellement

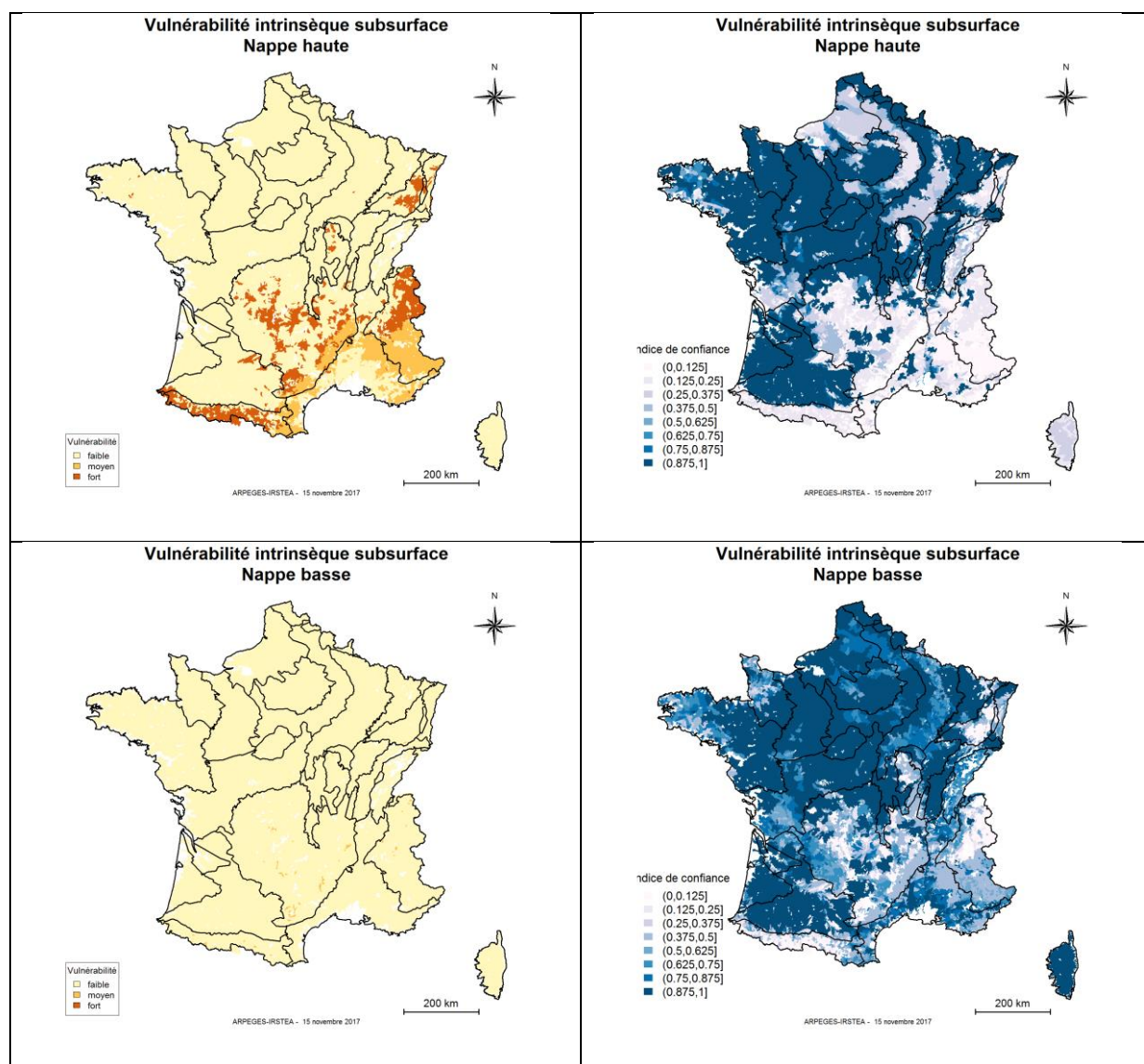


Figure 26 : Carte de vulnérabilité intrinsèque aux écoulements de subsurface

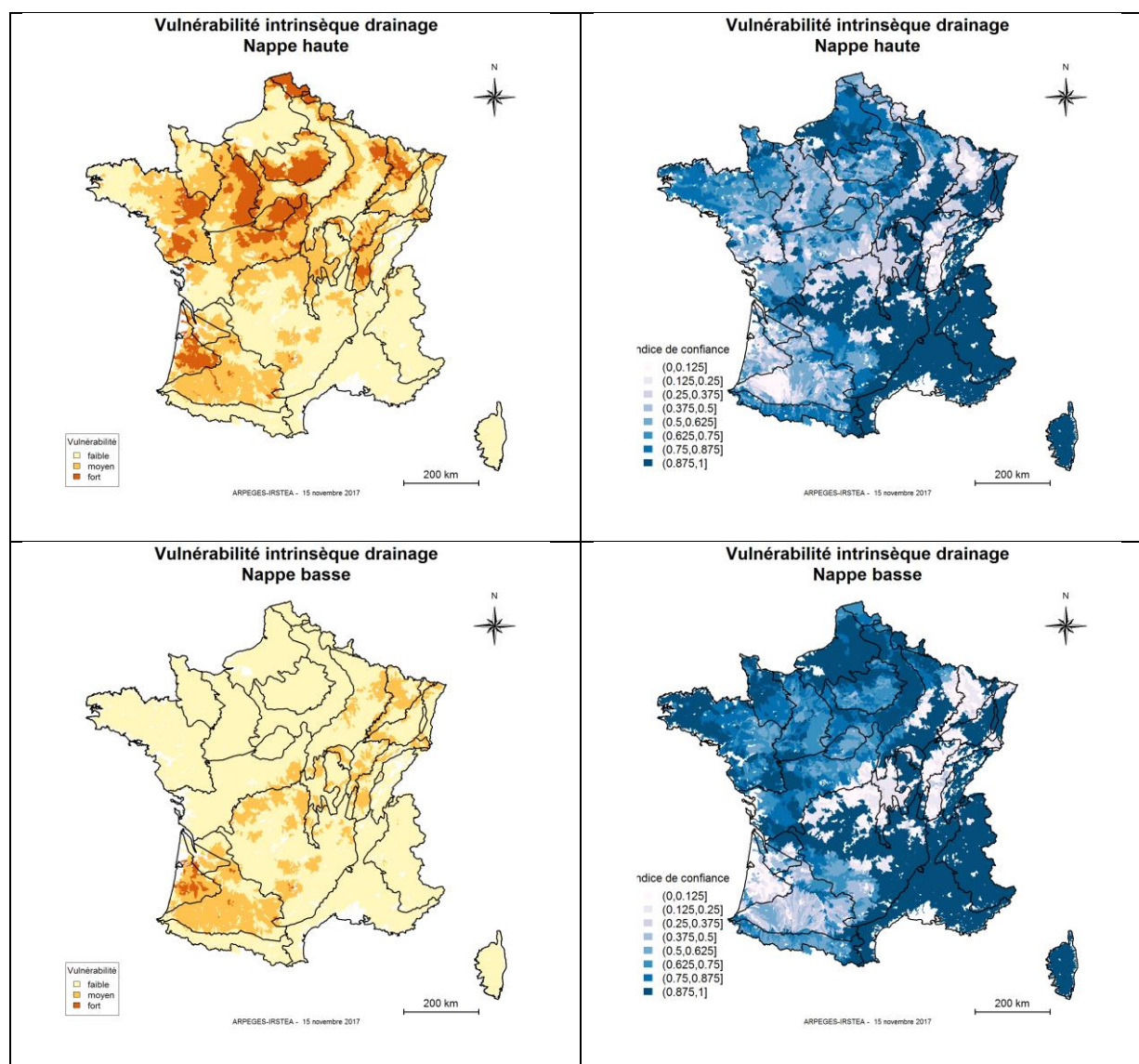


Figure 27 : Cartes de vulnérabilité intrinsèque au drainage agricole

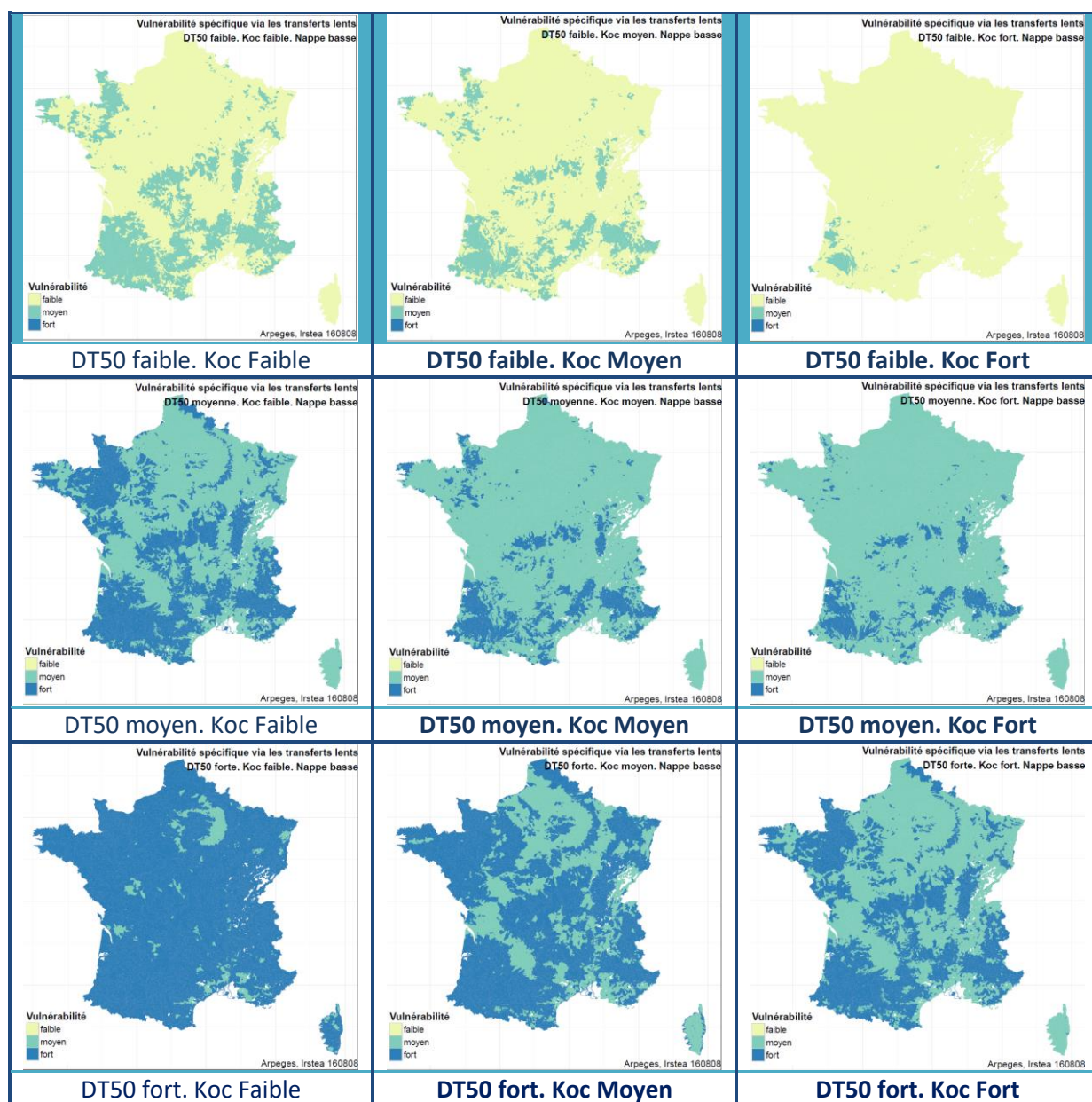


Figure 28 : Vulnérabilité spécifique aux transferts lents en période de Nappe Basse

DT ₅₀	Courte	Métaldéhyde		
	Moyenne	2,4-D 2,4-MCPA Aminotriazole Bentazone Nicosulfuron	Métazachlore S-métolachlore	
	Longue		Boscalid Chlortoluron Imidaclopride	Diflufénicanil Glyphosate Oxadiazon Pendiméthaline
		Faible	Moyen	Fort
		K _{oc}		

Tableau 13 : Répartition des quinze substances actives étudiées parmi les 9 combinaisons DT₅₀/K_{oc}

3.6. Application de la méthode via l'interface R

Les différentes parties ci-dessus décrivent le processus de construction qui a conduit à la version d'ARPEGES utilisée pour produire les résultats présentés dans ce rapport. Il s'agit d'une version basée sur les données d'entrée homogènes à toute la France les plus récentes, et sur les états et seuils de discrétisation les plus adaptés à ces données d'entrée.

Comme indiqué dans l'un des encadrés oranges, le cœur d'ARPEGES, constitué des tables de probabilités conditionnelles, n'est pas modifiable. Cependant, l'interface R laisse la main à l'utilisateur pour modifier une ou plusieurs données d'entrée, ainsi qu'un ou plusieurs seuils de discrétisation.

Le nombre d'étapes pour appliquer la méthode ARPEGES (à distinguer des étapes pour la construire) via l'interface R est donc variable selon les modifications apportées :

- ⇒ **Aucune modification, excepté le fait de travailler sur une nouvelle substance non étudiée dans le projet ARPEGES-EDL 2019 :**
 - 1^{ère} étape : récupération de la table des quantités supposées appliquées pour cette substance active (année(s) à choisir), répartition des quantités selon les deux saisons, agrégation des quantités à l'échelle des BVME, copié-collé de la donnée dans la table des données d'entrée.
 - 2^{ème} étape : lancement de l'interface R, en précisant « aucune modification ».
- ⇒ **Travail sur une nouvelle substance et modification de seuils de discrétisation pour une ou plusieurs variables :**
 - 1^{ère} étape : idem
 - 2^{ème} étape : lancement de l'interface R, en précisant « modification de seuil ».
- ⇒ **Travail sur une nouvelle substance et remplacement des données d'entrée pour une ou plusieurs variables :**

- 1^{ère} étape : idem
- 2^{ème} étape : récupération des nouvelles données d'entrée, agrégation des données à l'échelle des BVME, copié-collé de la ou les nouvelles donnée(s) dans la table des données d'entrée.
- 3^{ème} étape : choix de garder ou modifier les anciens seuils pour cette ou ces variables.
- 4^{ème} étape : lancement de l'interface R, en précisant « aucune modification » ou « modification de seuil » selon le choix fait en 3^{ème} étape.

Ce résumé vise seulement à faire visualiser les conséquences d'un choix de modification de données d'entrée et/ou des seuils de discrétisation. Pour plus de précision sur la mise en œuvre de l'interface, se référer au guide utilisateur dédié.

4. Méthode de comparaison des sorties ARPEGES à des données observées

4.1. Traitement des données observées comparées : méthode d'intégration des courbes enveloppes (sur 3 ans)

Dans le cadre de travaux précédents, Irstea a développé une méthode d'interprétation des données de surveillance, visant à dépasser les limites liées à l'hétérogénéité des caractéristiques de suivi (limites de quantification, fréquence des analyses, répartition des analyses par rapport aux événements hydrologiques). Ainsi, l'analyse station par station n'est pas satisfaisante pour évaluer des tendances et dynamiques de la contamination pour une substance.

La méthode visant à pallier ces difficultés d'analyse consiste à regrouper les observations/concentrations mesurées par zone géographique et par période de quinze jours, et à calculer le 90^{ème} centile de chaque groupe de données. Il ne s'agit plus d'une analyse spécifique d'une station donnée, mais de l'étude de la courbe enveloppe de la contamination de la zone considérée.

Le découpage géographique retenu est le résultat de deux étapes : la 1^{ère} étape correspond aux Hydro-Eco-Régions (HER) de niveau 2 telles que définies par Wasson et al (2002), basées sur la géologie, le relief et le climat (facteurs environnementaux importants dans le transfert de substances actives) ; la 2^{ème} étape est le regroupement des HER de niveau 2 adjacentes en fonction des Orientations Technico-Economiques des Exploitations (OTEX) majoritaires des zones (Figure 30), afin d'avoir un découpage assez homogène d'un point de vue de l'occupation des sols, tout en ayant des zones suffisamment vastes pour mener des analyses statistiques raisonnables. Le résultat du découpage est présenté en Figure 31.

Les données observées utilisées pour tester la cohérence des sorties d'ARPEGES sont les données des réseaux du Programme de Surveillance des Agences de l'Eau, centralisées par le SDeS (Agences de l'Eau, Sandre, 2017). Ces données sont synthétisées sous la forme d'un indicateur spatial de la contamination par HER modifiée. Il s'agit de la valeur moyenne de la courbe enveloppe (90^{ème} centile par quinzaine de jours) de la concentration dans le milieu sur la **période 2013-2015**, calculée en rapportant l'aire sous la courbe à la durée sur laquelle elle peut être calculée, exprimée en nombre de quinzaines de jours pour lesquelles on a une valeur observée. La valeur de la moyenne calculée, assimilable à une concentration moyenne (relative au 90^{ème} centile) est ainsi associée au nombre de quinzaines (au sens de moitié de mois) sur lequel elle porte (de 1 à 72), ce qui permet de qualifier la

« confiance » en terme de représentativité que l'on peut accorder à la valeur de concentration moyenne calculée, compte tenu du nombre de données disponibles. La période d'intégration de 3 ans choisie vise à lisser les variabilités interannuelles des concentrations.

Cette intégration moyennée par le nombre de quinzaines est illustrée en Figure 29. Il faut d'abord noter que l'intégration de courbe enveloppe se fait sur toute la durée de la période 2013-2015⁹, et non sur certaines périodes de transfert. Afin d'assurer une continuité des intégrales entre plusieurs années, un rectangle correspondant à la largeur d'une semaine est ajouté à gauche du point de la première quinzaine, et à droite du point de la dernière quinzaine. De même, afin de compter les points isolés ou en bord d'une série de points, un rectangle de durée d'une quinzaine ou d'une semaine est ajouté selon les cas.

Enfin, les mesures inférieures à la limite de quantification sont prises en compte dans les courbes enveloppes et leurs intégrales, en leur attribuant la valeur de la moitié de la limite de quantification.

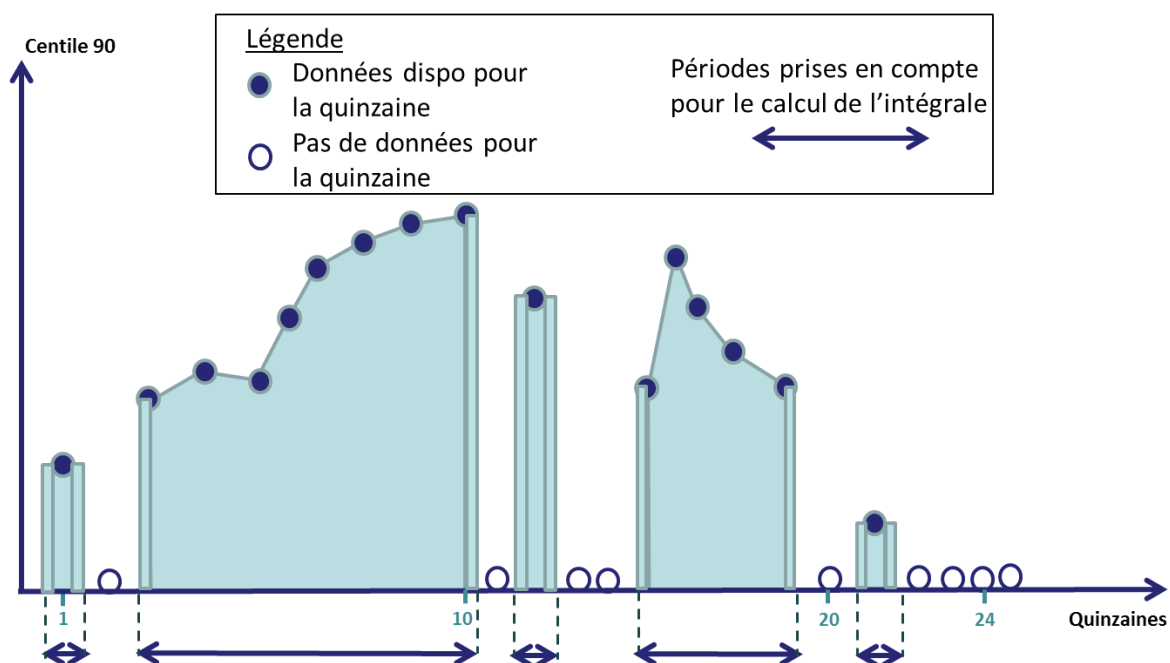


Figure 29 : Illustration de la méthode de calcul de l'intégrale "moyennée" par le nombre de quinzaines

⁹ La période d'intégration correspond aux trois années civiles, et non aux trois années hydrologiques de la période 2013-2015. On considère ici que le lissage lié au fait d'intégrer sur trois ans permet de ne pas se baser sur les limites d'années hydrologiques pour les substances transférant en période hivernale.

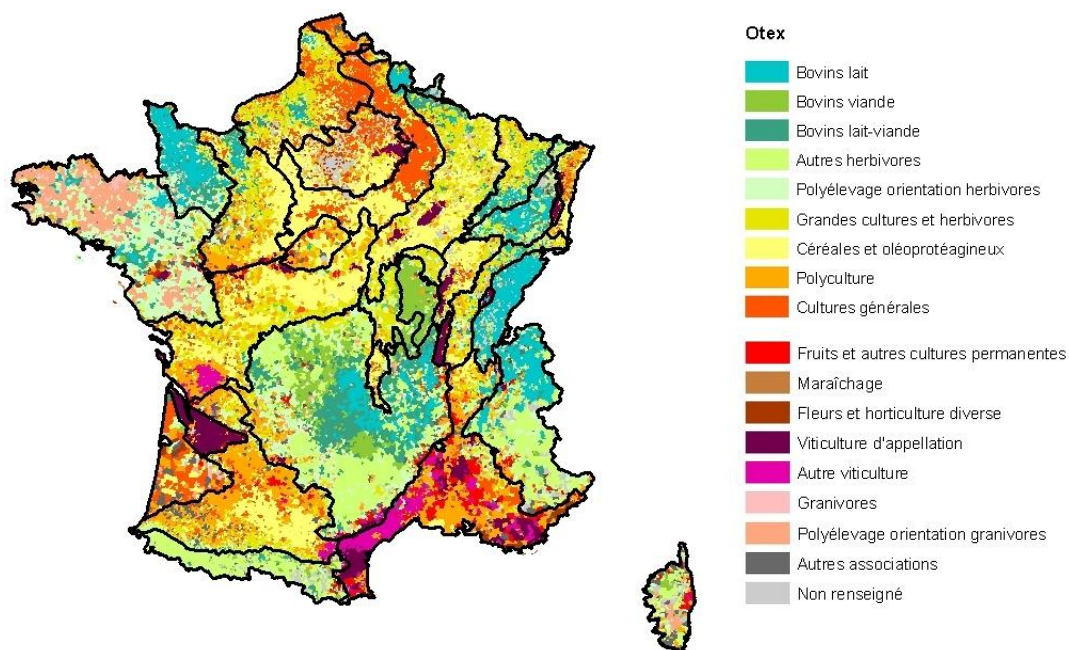


Figure 30 : Découpage national en HER modifiées par la prise en compte de l'occupation du sol (OTEX)

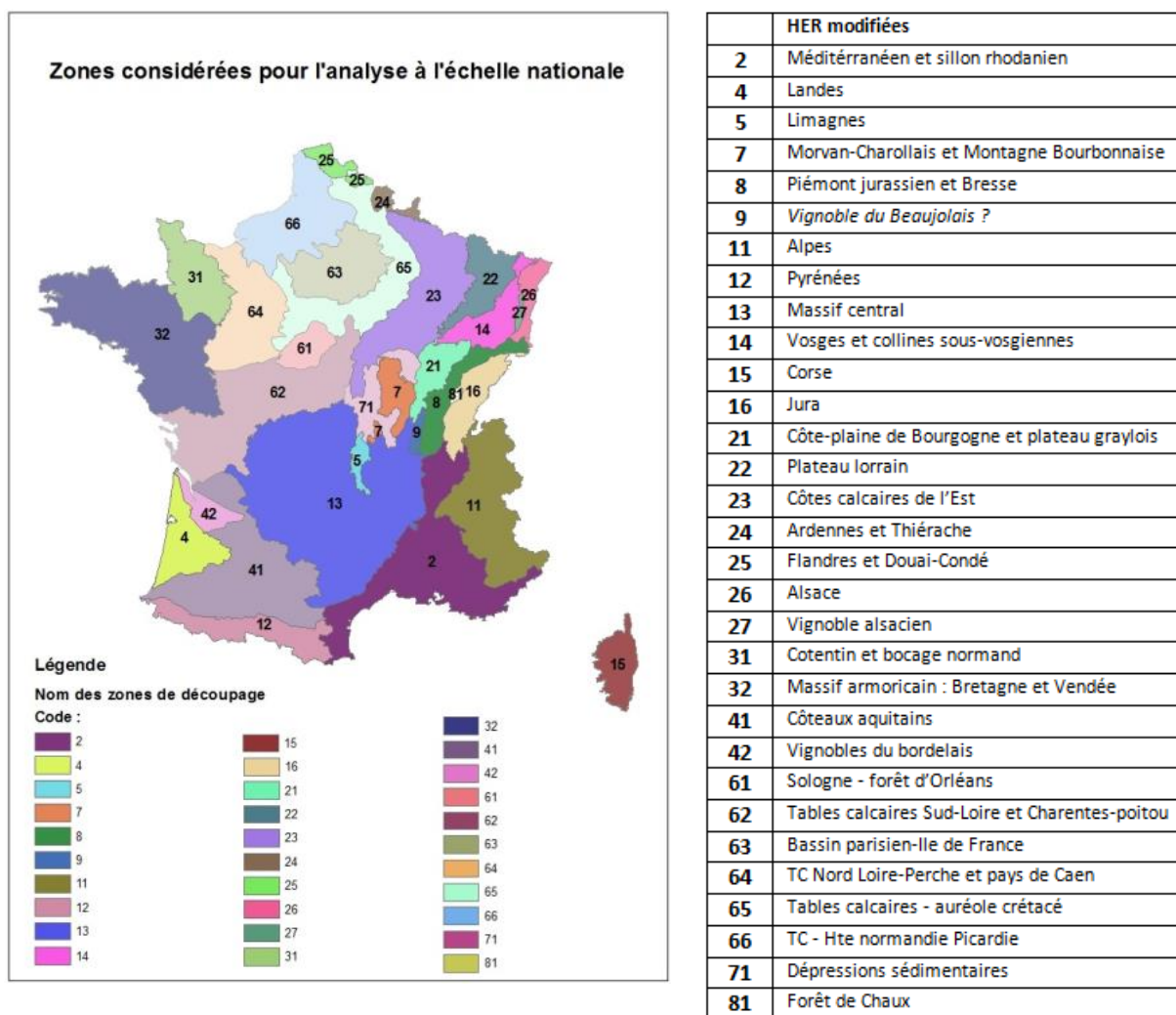


Figure 31 : Position et dénomination des HER modifiées par la prise en compte de l'occupation du sol (OTEX)

4.2. Sorties ARPEGES comparées

Les sorties ARPEGES les plus appropriées pour le test de cohérence du modèle sont les cartes de potentiel de contamination via les transferts lents. En effet, l'intégrale « moyennée » des données observées s'apparente à la résultante des transferts lents sur 3 ans. Une seule saison sur les deux est comparée aux données observées, et dépend de la substance active : il s'agit de la saison la plus pertinente pour chaque substance active, c'est-à-dire celle dont le coefficient de saisonnalité est le plus grand.

On entrevoit déjà des différences entre les deux éléments comparés :

- Des échelles de calcul très différentes : BVME pour ARPEGES, HER modifiée pour les données observées.
- Des échelles de temps différentes : les données de pression d'usage phytosanitaire (achat des produits) datent de 2015 dans ARPEGES, tandis que les données mesurées sont intégrées sur 2013-2015.

Cependant, on insiste ici sur le fait qu'il ne s'agit pas d'une validation de modèle, mais bien d'une comparaison pour tester la cohérence globale des sorties d'ARPEGES.

Les seuils des classes de 90^{ème} centile moyen des données observées sont identiques pour toutes les substances actives et ont été ajustés de façon à obtenir la meilleure cohérence entre les deux types de cartes pour l'ensemble des substances actives.

5. Test de cohérence du modèle

5.1. Résultats de la comparaison et analyse

Les comparaisons entre les cartes ARPEGES et les cartes de données observées sont présentées en

Figure 32, Figure 33 et Figure 34. Il ressort de cette comparaison une cohérence globale satisfaisante entre ARPEGES et les données observées : en considérant l'échelle France, la hiérarchisation des zones des classes ARPEGES est globalement similaire à celle des classes de données mesurées.

Bien qu'ayant été ajustées au mieux, il y a pour certaines substances un décalage d'une classe entre les cartes ARPEGES et les cartes de données mesurées, mais le fait que la hiérarchisation soit la même est déjà satisfaisant.

Comme attendu, si l'on regarde plus en détail au sein des HER, on constate des différences, tout du moins il est difficile de trancher pour une HER données sur la cohérence entre les deux cartes. Ceci vient de la différence importante de maille de calcul entre les deux éléments comparés.

Les classes de 90^{ème} centile moyen ajustées aux classes ARPEGES sont [0 ; 0,005], [0,005 ; 0,012], [0,012 ; 0,05], [0,05 ; 1,07], en µg/L. Les seuils de classes sont très faibles, les deux premières classes contiennent des valeurs inférieures ou proches de la limite de quantification moyenne en France (0,01 µg/L). On rappelle ici que l'intégrale de la courbe enveloppe des concentrations est réalisée sur toute la période des 3 années, et non uniquement sur les périodes de transfert. Les faibles concentrations des périodes hors-transfert font donc baisser fortement les 90^{èmes} centiles moyens calculés : il n'est donc pas pertinent de comparer ces valeurs à des Normes de Qualité Environnementale (NQE), qui font référence à des concentrations en période de transfert.

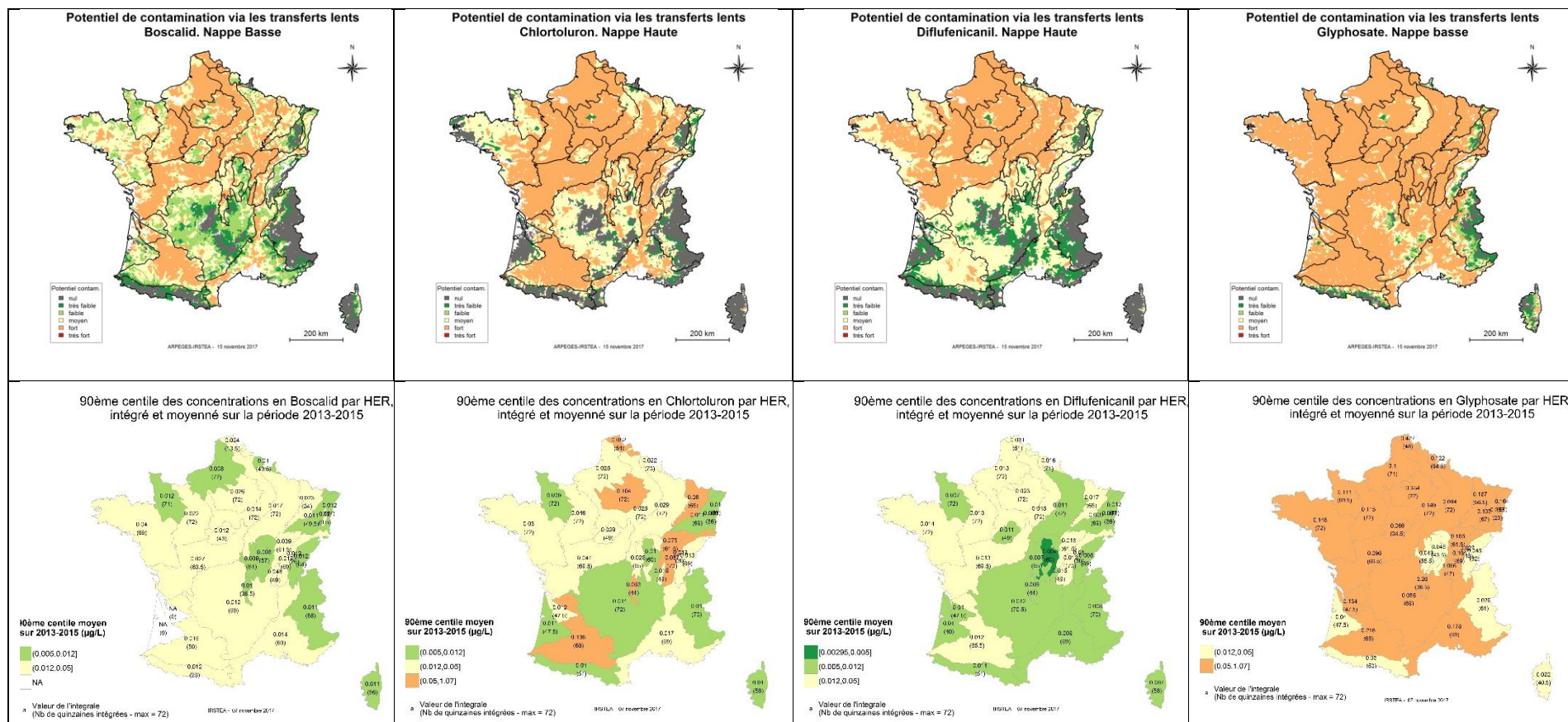


Figure 33 : Comparaison des cartes ARPEGES et de données observées pour le Boscalid, le Chlortoluron, le Diflufenicanil et le Glyphosate

5.2. Cas particuliers incohérents

Malgré cette cohérence globale relevée dans le paragraphe précédent, on constate deux cas extrêmes qui ne supportent pas la conclusion générale :

L'aminotriazole (

- Figure 32) : le potentiel de contamination semble globalement bien plus faible que les 90^{èmes} centiles moyens. On constate pour cette substance que 21 % des quantités totales achetées sont considérées comme non-agricoles par la méthode de spatialisation de l'INRA, ne sont pas spatialisées (80 870 kg) et ne sont donc pas prises en compte pour cette étude. Une partie des cibles de l'aminotriazole s'intitule « traitements généraux ». Dans la méthode INRA, la cible « traitements généraux » est totalement considérée comme étant sur des surfaces non-agricoles (excepté le glyphosate). Ainsi, ces quantités d'aminotriazole considérées comme appliquées sur non-agricole pourraient venir pour tout ou partie de ces traitements généraux, et expliquer les écarts entre les résultats d'ARPEGES et les valeurs mesurées.
- La pendiméthaline (Figure 34) : il s'agit du cas inverse de l'aminotriazole. Le potentiel de contamination est fort sur une majeure partie des surfaces françaises, tandis que les 90^{ème} centiles moyens indiquent une contamination observée faible voire très faible. Une explication possible de cette grande différence est le Koc très élevé (15740 mL.g⁻¹) de cette substance. Il est possible que, dans ce cas, on dépasse le domaine de validité d'ARPEGES, et que, bien qu'elles soient largement utilisées, le modèle surestime le transfert pour ces substances s'adsorbant très fortement.

6. Etude de cas et pistes d'interprétations

Les façons d'analyser les résultats ARPEGES peuvent être multiples :

- Identification des substances actives les plus à risque sur une zone donnée ;
- Identification des masses d'eau les plus à risque pour une substance donnée ;
- Identification des saisons (nappe haute ou nappe basse) et types de contamination (transferts lents ou rapides) les plus à risque selon les zones et substances ;
- Identification des causes majeures de contamination selon les zones.

L'objectif de cette partie n'est pas de fournir une analyse détaillée des résultats pour chaque substance selon ces différentes façons d'analyser, mais de porter l'attention sur des éléments importants à prendre en compte pour toute analyse (indice de confiance et probabilité de chaque état d'une variable), en prenant l'exemple de la démarche d'identification des causes majoritaires de contamination.

Le cas du S-métolachlore en nappe basse et transferts lents sera pris pour illustrer la majeure partie du propos, ainsi que celui du boscalid pour un exemple précis.

6.1. Etude du potentiel de contamination

L'étude du potentiel de contamination pour remonter aux causes majoritaires du transfert commence par l'identification du zonage de chaque classe de potentiel. Déjà à ce stade, il est important de tenir compte de la carte d'indice de confiance, qui permet de nuancer les résultats et distinguer des « sous-zones » au sein d'une classe de potentiel.

Par exemple, sur la Figure 35 (potentiel de contamination du S-métolachlore en nappe basse via les transferts lents et son indice de confiance), on remarque que le potentiel de contamination fort est obtenu :

- avec un indice de confiance très faible compris entre 0 et 0,125 pour les BVME de la HER 61 (Sologne, cf Figure 31) et à l'ouest de la HER 32 (Bretagne) ;
- avec un indice de confiance compris entre 0,125 et 0,25 pour les BVME de la HER 31 (Cotentin et bocage normand) ;
- et avec un indice de confiance compris entre 0,25 et 0,375 pour les BVME du sud-ouest de la HER 41 (coteaux aquitains).

Pour rappel, un indice de confiance proche de 0 signifie que l'écart entre les deux plus grandes probabilités est très faible (cf 3.5.1.3). Pour un indice de confiance nul, un cas de figure extrême est d'avoir les probabilités des cinq classes à 20% ; un autre cas extrême est d'avoir deux classes chacune avec une probabilité de 50%. Pour avoir plus d'informations, il est donc utile d'étudier les cartes de probabilités d'une ou plusieurs classes de potentiel de contamination.

Par exemple, pour les BVME de la HER 62 (tables calcaires Sud-Loire et Charente-Poitou), la probabilité de la classe très fort est homogène entre les BVME, comprise entre 0 et 0,125, et la somme des probabilités des classes fort et très fort est quasiment homogène, comprise entre 0,5 et 0,625. Par ailleurs, on constate pour ces mêmes BVME que le potentiel de contamination (décision finale) est soit moyen, soit fort, alors que l'indice de confiance est pour tous compris entre 0 et 0,125. Il est donc probable que les classes « moyen » et « fort » soient très proches en termes de probabilités pour les BVME de cette HER (il faudrait une analyse de la carte de probabilité de la classe moyen pour le confirmer), et que l'on ait ici un effet de seuil en prenant la probabilité majoritaire pour la carte de décision finale. Cet exemple a pour but d'attirer l'attention sur l'importance de relativiser les résultats de la carte de décision finale du potentiel de contamination pour en saisir toutes les nuances.

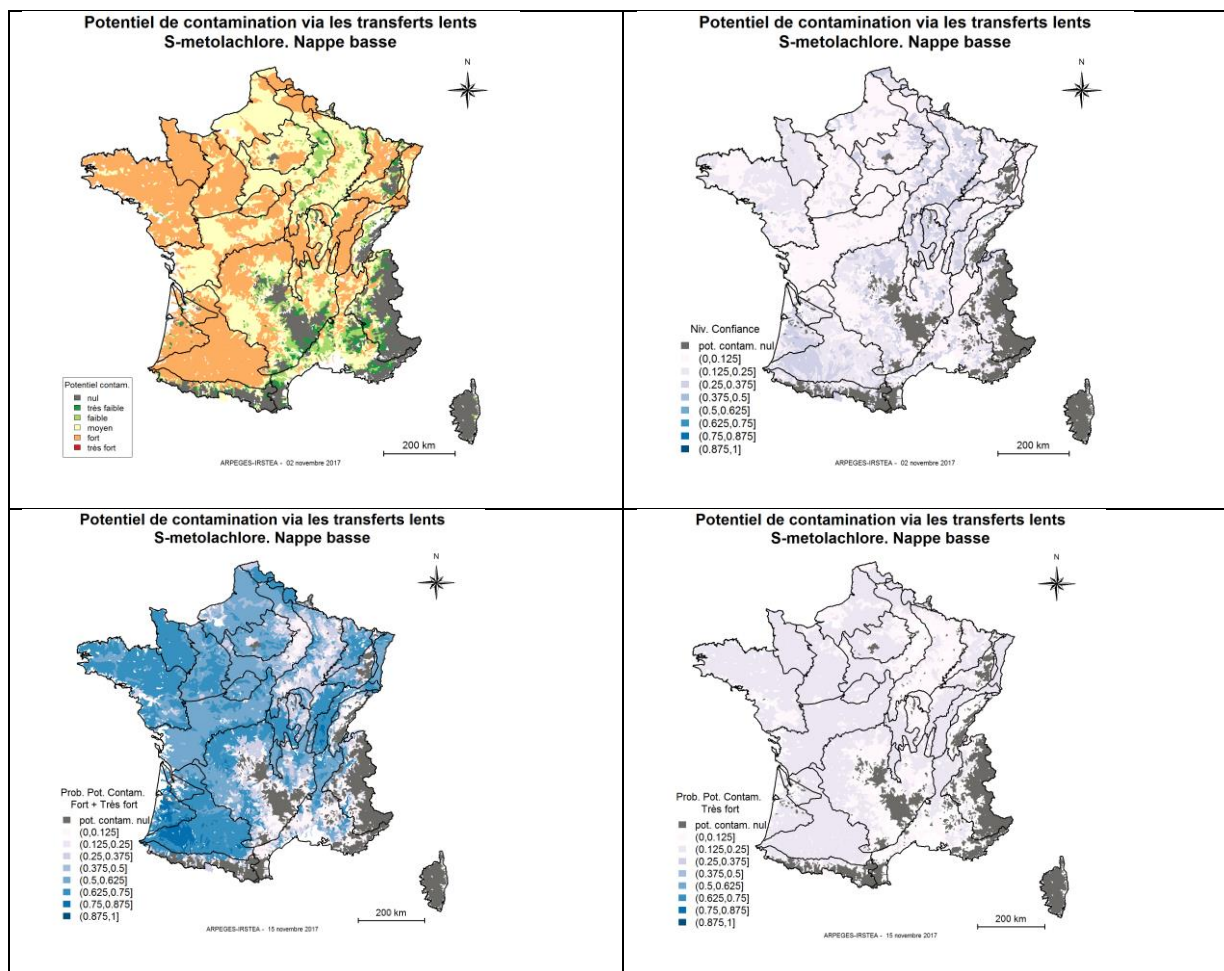


Figure 35 : Analyse conjointe des cartes du potentiel de contamination, de son indice de confiance, de la somme des probabilités des classes fort et très fort, et de la probabilité de la classe très fort

6.2. Causes du potentiel de contamination

Après avoir analysé le potentiel de contamination, l'analyse conjointe de la pression d'usage phytosanitaire et de la vulnérabilité spécifique permet d'apporter des premiers éléments sur les causes du potentiel de contamination. Là encore, il est important de considérer tous les résultats de la vulnérabilité spécifique (décision finale, indice de confiance et probabilité des classes). Pour la pression d'usage, il n'y a pas de carte d'indice de confiance ni de carte de probabilité des classes, car il s'agit d'une donnée d'entrée pour laquelle la valeur de pression sur le BVME est traduite par une probabilité de 100 % pour la classe correspondante (cf Tableau 3).

L'analyse des causes du potentiel de contamination qui suit s'intéresse d'abord aux causes du potentiel de contamination fort, puis aux causes du potentiel de contamination moyen.

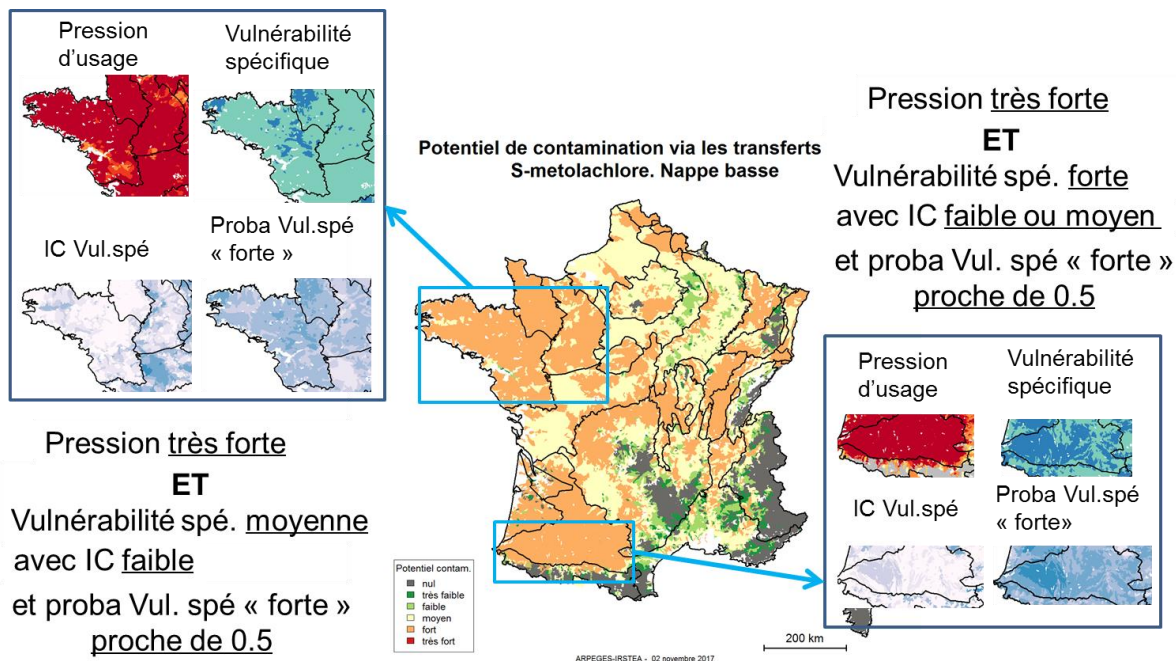
Pour le potentiel de contamination « fort », deux cas de figure sont rencontrés (

Figure 36 et Tableau 14). Il en ressort que le potentiel de contamination fort est toujours associé à une pression d'usage forte. En ce qui concerne la vulnérabilité spécifique, elle est, dans les deux

exemples, le plus souvent tangente entre les classes « moyen » et « fort », sauf dans des cas particuliers (« triangle » à l'ouest de l'encadré coteaux aquitains), pour lesquels la vulnérabilité spécifique est dans la classe « fort » avec un indice de confiance plus élevé. Dans cette zone, on peut conclure qu'à la fois les usages et la vulnérabilité spécifiques contribuent au potentiel de contamination « fort » et que par conséquent, il serait intéressant d'agir à la fois sur ces deux facteurs pour le limiter. Il y a notamment potentiellement plus d'intérêt à agir sur la vulnérabilité spécifique sur cette zone que sur les autres.

Localisation de l'encadré	Potentiel de contamination	Pression d'usage	Vulnérabilité spécifique (décision finale)	Commentaires
Bretagne	Fort	Très fort	Moyen	La vulnérabilité spécifique représentée est en classe « moyen » (probabilité majoritaire), mais, d'une part, l'indice de confiance associé est faible et, d'autre part, la probabilité de la classe de vulnérabilité spécifique « fort » est proche de 0,5. On est donc dans un cas de vulnérabilité spécifique « tangent » entre classes « moyen » et « fort »
Sud de l'HER 41 (coteaux aquitains)	Fort	Très fort	Fort	Pour la partie Est de cette zone, on est dans un cas similaire à celui présenté ci-dessus : les deux classes de vulnérabilité spécifique « moyen » et « fort » représentées sur la carte finale sont proches. En revanche, sur la partie Ouest, l'indice de confiance de la vulnérabilité spécifique « fort » est un peu plus élevé indiquant que la vulnérabilité spécifique a ici plus d'influence sur le potentiel de contamination « fort » calculé.

Tableau 14 : Analyse des deux cas de figure pour le potentiel de contamination « fort » pour le S-métolachlore



Légende des encadrés :

Pression d'usage	Vulnérabilité spécifique	IC Vul. Spé et Proba Vul. Spé « forte »
<ul style="list-style-type: none"> nulle faible moyen fort 	<ul style="list-style-type: none"> faible moyen fort 	<ul style="list-style-type: none"> (0,0,0.125] (0.125,0.25] (0.25,0.375] (0.375,0.5] (0.5,0.625] (0.625,0.75] (0.75,0.875] (0.875,1]

Figure 36 : Analyse des causes du potentiel de contamination « fort » sur deux exemples (Bretagne à gauche et Coteaux aquitains à droite)

Les deux exemples suivants concernent l'interprétation d'un potentiel de contamination «moyen» et un potentiel de contamination variant de « faible » à « fort » (Figure 37, Figure 38 et Tableau 15). Dans le premier cas, sur l'exemple du S-métolachlore, la vulnérabilité spécifique « moyen » atténue l'influence de la pression d'usage « fort » dans la note finale de potentiel de contamination « moyen », tandis que dans le second cas, sur l'exemple du Boscalid, c'est la pression d'usage (variant de « faible » à « fort ») qui module l'influence de la vulnérabilité spécifique « fort » dans la valeur finale du potentiel de contamination.

Substance active	Potentiel de contamination	Pression d'usage	Vulnérabilité spécifique (décision finale)	Commentaires
S-métolachlore	Moyen	Très fort	Moyen	L'indice de confiance de la vulnérabilité spécifique « moyen » est élevé, tandis que la probabilité de la classe « fort » de vulnérabilité spécifique est très

				faible. La vulnérabilité spécifique « moyen » est donc confortée et contribue à atténuer la valeur finale « moyen » du potentiel de contamination en dépit d'une pression phytosanitaire très forte.
Boscalid	Moyen	Moyen	Fort	La vulnérabilité spécifique « fort » est confortée par un indice de confiance « moyen » et une probabilité de cette classe « fort » élevée. Elle contribue donc au potentiel de contamination mais, contrairement au cas précédent, c'est la pression d'usage qui module la valeur finale de ce dernier.

Tableau 15: Analyse des deux cas de figure pour le potentiel de contamination « moyen » pour le S-métolachlore et le boscalid

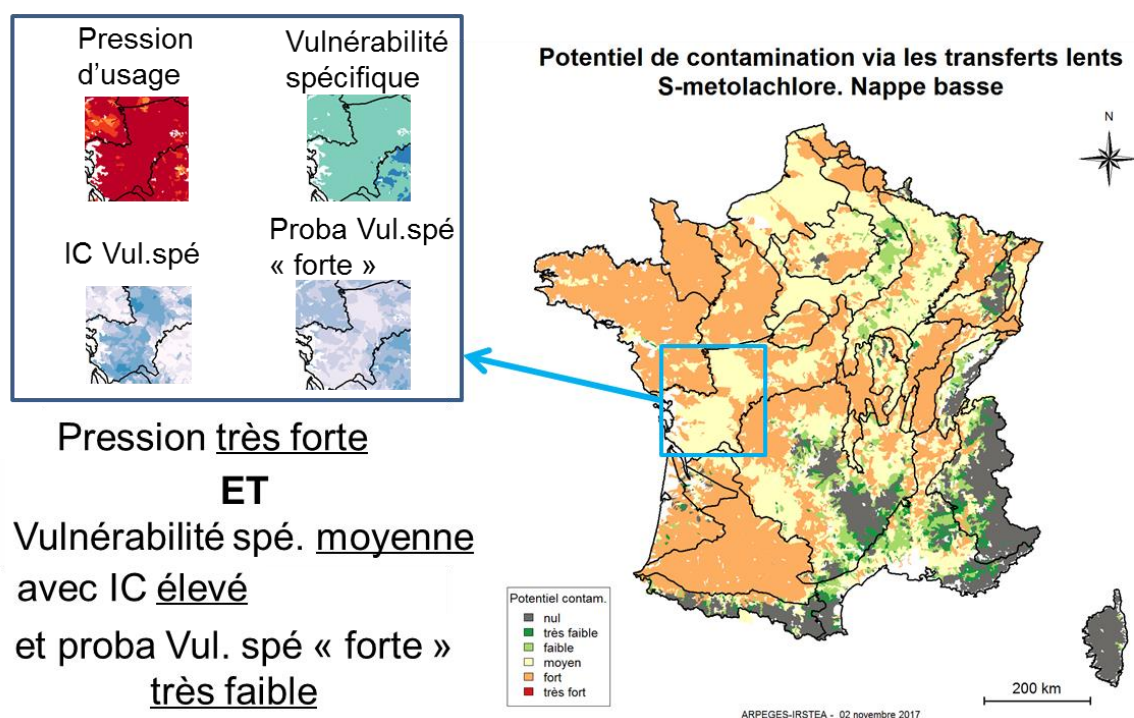


Figure 37 : Analyse des causes du potentiel de contamination « moyen » du S-métolachlore (légende de l'encadré en Figure 36)

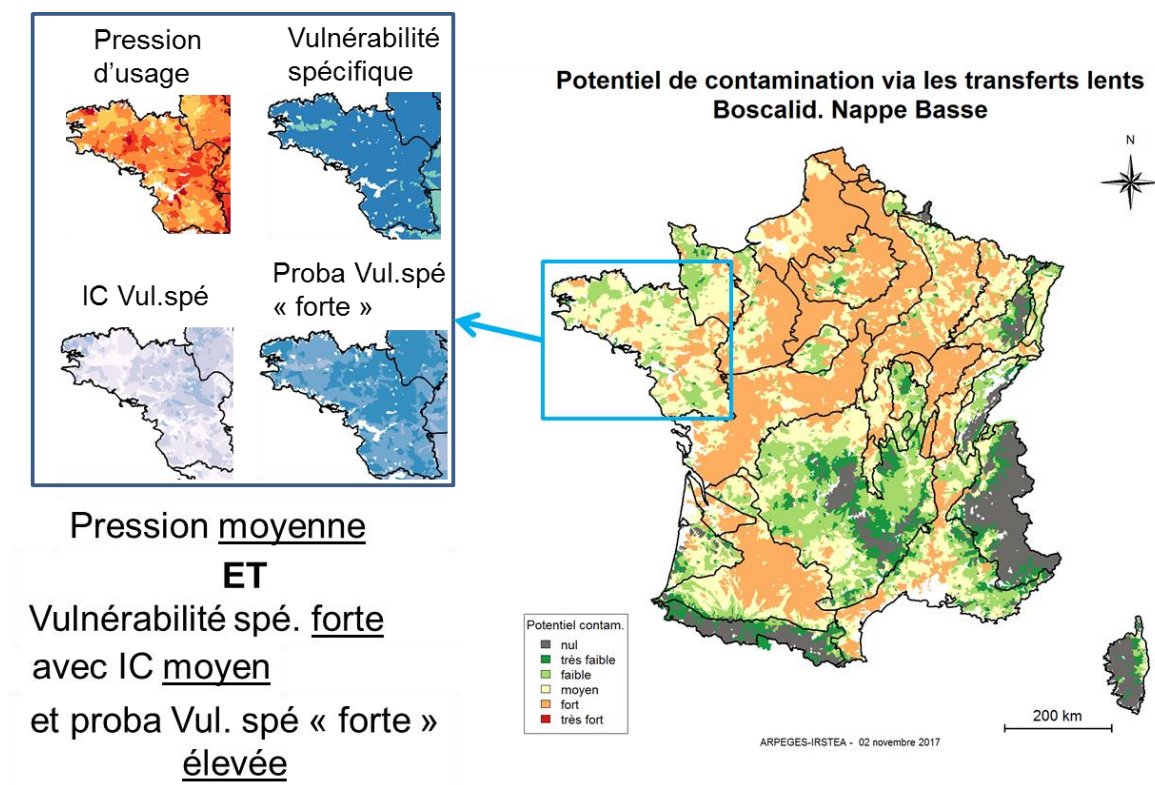


Figure 38 : Analyse des causes du potentiel de contamination « moyen » du Boscalid (légende de l'encadré en Figure 36)

6.3. Causes de la vulnérabilité spécifique

Après avoir constaté une vulnérabilité spécifique tangente entre les classes « moyen » et « fort » pour le S-métolachlore en nappe basse dans les zones Bretagne et Coteaux aquitains, il est intéressant de creuser quels sont les types d'écoulement principalement en cause selon les zones.

Sur la Figure 39, on constate que la vulnérabilité spécifique aux transferts lents en Bretagne est uniquement due à la vulnérabilité intrinsèque au ruissellement, tandis que la même vulnérabilité spécifique en Coteaux aquitains est à la fois due aux vulnérabilités intrinsèques au ruissellement et au drainage agricole. Il faudrait à nouveau étudier les indices de confiance associés pour confirmer ces observations (non présenté ici).

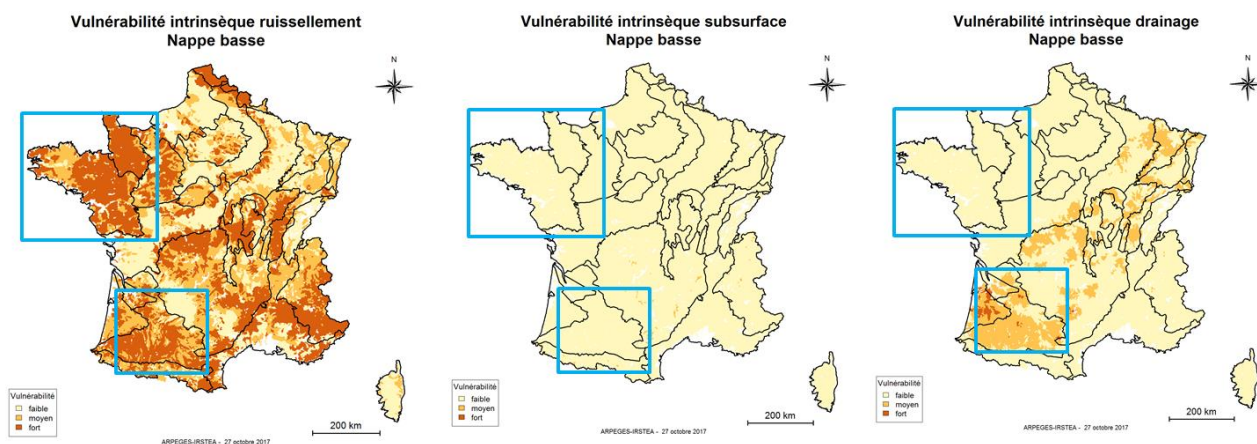


Figure 39 : Comparaison de la vulnérabilité intrinsèque des trois types d'écoulement en les deux zones "focus"

Pour remonter aux causes majoritaires de ces vulnérabilités intrinsèques, il faudrait ensuite étudier une à une les données d'entrée (cf 3.4 et annexes correspondantes) afin de mieux cerner les leviers d'action et zones les plus pertinentes pour les mettre en œuvre (par exemple agir sur le ruissellement pour limiter la vulnérabilité intrinsèque sur la zone Ouest des coteaux aquitains).

6.4. Liens entre potentiels de contamination via transferts lents ou rapides

Le même raisonnement que précédemment peut être appliqué pour les cartes de potentiel de contamination via les transferts rapides.

Il est également intéressant de comparer entre elles les cartes de potentiel de contamination via les transferts lents et rapides. Par exemple, pour le cas du S-métolachlore (Figure 40), on remarque que le potentiel de contamination via les transferts rapides est partout plus faible que celui via les transferts lents, sauf au centre de la HER 31 (Cotentin et bocage normand), de la HER 4 (Landes) et au sud-ouest de la HER 41 (coteaux aquitains), où il reste fort. Cela incite donc dans ce cas à prioriser les actions à l'échelle nationale sur ces zones où le potentiel de contamination est fort à la fois via les transferts lents et rapides. Cela permet également d'identifier les zones où il serait plus intéressant d'agir en priorité pour limiter la vulnérabilité intrinsèque liée aux transferts rapides (aménagement de zones tampons, pratiques de limitation du ruissellement, ...).

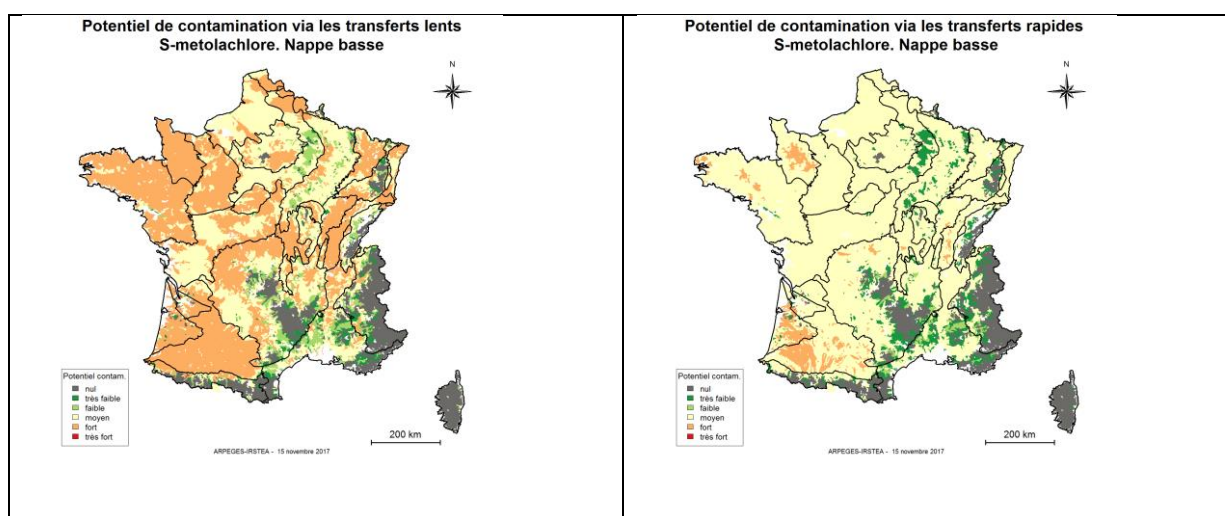


Figure 40 : Comparaison des cartes de potentiel de contamination via les transferts lents (à gauche) ou rapides (à droite) pour le S-métolachlore

7. Conclusion

7.1 La méthode ARPEGES et son application au niveau national

La méthode ARPEGES (Analyse de Risque Pesticides pour la Gestion des Eaux de Surface) a pour objectifs d'évaluer le potentiel de contamination des masses d'eau de surface par les produits phytosanitaires. Elle rend compte des connaissances expertes mobilisables sur le devenir des produits phytosanitaires dans l'environnement et leur transfert hydrique vers les eaux de surface, en fonction des trois grands ensembles de facteurs déterminant le devenir et le transfert hydrique d'une substance active dans le milieu : pour chaque substance active, ARPEGES permet ainsi de **croiser la vulnérabilité du milieu** au transfert des produits phytosanitaires, les **caractéristiques propres de la molécule**, et la **pression d'usage exercée par cette substance** sur le territoire français métropolitain.

Cette méthode est basée sur la construction d'un modèle de connaissance des processus de transfert. Elle distingue les processus se produisant en période de nappe haute et en période de nappe basse. La méthode a été développée, à l'origine, à l'échelle nationale. A cette échelle, les données sont hétérogènes, et lacunaires et les processus complexes : afin de fournir une indication du niveau de confiance associé au résultat, le **croisement** des données se fait de façon **probabiliste** à l'aide d'un **système bayésien**. Au total, 18 variables sont agrégées pour obtenir le potentiel de contamination. L'échelle d'étude est le **Bassin Versant de Masse d'Eau (BVME)** défini dans le cadre de la DCE (environ 10 000 BVME sur le territoire métropolitain).

Dans le cadre de l'élaboration de l'état des lieux 2019 de la Directive Cadre sur l'Eau, la méthode a été appliquée à l'échelle nationale pour 15 substances actives prioritaires ou d'intérêt identifiées par les Agences de l'Eau, en utilisant les données d'usage décrit par les achats de l'année 2015 issues de la BNVD, spatialisées à partir des données RPG 2014. Par rapport aux applications précédentes d'ARPEGES, la méthode a été cette fois appliquée pour une substance active seule, et en utilisant des seuils identiques pour l'ensemble des molécules pour caractériser les niveaux de pression d'usages (5 classes de très faible à très fort).

Les résultats obtenus ont été confrontés aux données de surveillance Eaux de Surface. Ces données ont été agrégées par Hydro-Eco-Région (HER) pour assurer leur représentativité, et moyennées sur les années 2013-2016, afin de lisser la variabilité interannuelle : cette différence d'échelle spatiale et temporelle doit être gardée à l'esprit dans l'interprétation des résultats obtenus. Dans l'ensemble, les résultats de potentiel de contamination par transferts lents (en nappe haute ou en nappe basse selon la période d'application principale de la substance) prédits par ARPEGES montrent une relative cohérence avec les niveaux de concentrations (concentrations moyennes annuelles de la DCE) observées. Compte tenu des différences d'échelle de restitution, des incertitudes sur les données, notamment de pression d'usage, et de la nature probabiliste des résultats fournis par ARPEGES, cette confrontation doit toutefois être effectuée avec précaution. Ce rapport s'attache à donner des pistes d'interprétation des résultats, en illustrant notamment comment les différentes cartes de résultat : potentiel de contamination majoritaire, niveau de confiance et probabilité de potentiel de contamination fort et très fort, gagnent à être exploitées en parallèle.

L'un des avantages essentiels de la méthode ARPEGES est de permettre de remonter aux causes des potentiels de contamination élevée, en faisant la part des choses entre la pression d'usage de la

molécule, ses caractéristiques propres (DT50 et Koc) et la vulnérabilité du milieu aux différents types de transfert (ruissellement, drainage, écoulements de subsurface). Cela permet d'identifier les leviers sur lesquels agir en priorité : diminution des usages, substitution, aménagement des bassins versants (pour limiter le ruissellement par exemple).

7.2 L'interface ARPEGES et ses fonctionnalités pour une application locale

En parallèle à l'application d'ARPEGES au niveau national, telle que décrite dans ce rapport, une interface ARPEGES a été développée pour en permettre une utilisation autonome (Mellac et Piffady, 2018), éventuellement réduite à un territoire infra-national (typiquement le territoire d'action de chaque Agence). Cet outil permet aux Agences d'appliquer ARPEGES pour d'autres substances que celles étudiées dans ce rapport. Il permet également d'adapter des seuils de discrétisation des variables d'entrée pour discriminer des variables d'une façon plus pertinente au niveau local que ne l'autorise l'application réalisée au niveau national. Cette adaptation doit toutefois être menée dans le cadre de la méthode ARPEGES telle que définie au niveau national : le nombre de classes de discrétisation pour les données d'entrée ne peut être modifié, et les BVME (nombre, limites, identifiant) sont ceux qui ont été utilisés dans l'application au niveau national. Pour le premier point, une évolution demanderait de reconstruire les tables de probabilité du modèle bayésien sur lequel est basé ARPEGES, ce qui est actuellement inenvisageable ; pour le deuxième point, modifier les BVME supposerait de recalculer l'ensemble des variables d'entrée pour chacun des BVME : ce travail est réalisable, mais demande de mettre en œuvre différentes techniques d'agrégation/désagrégation des données (cf section 3.4.1). Il semblerait judicieux que ce travail soit mené au niveau national, ou au moins facilité par la mise à disposition de routines dédiées par l'AFB.

Quelques types de données ont d'ores et déjà été identifiées comme méritant d'être améliorées, selon le cas au niveau national ou régional :

- **données relatives aux usages.** Il serait souhaitable de disposer des quantités appliquées à une maille spatiale plus fine qu'actuellement, de limiter les données orphelines (c'est-à-dire non spatialisées) auxquelles mène la méthode actuelle de spatialisation des données d'achat de la BNVD (projet Ecophyto - INERIS, INRA/ODR), de définir des coefficients de répartition saisonnière plus représentatifs des usages par bassin. Prendre en compte la spatialisation des parcelles en agriculture biologique (AB) paraîtrait également souhaitable, et ce d'autant plus que les surfaces en AB vont croissant,
- **carte de la battance** plus précise sur les bassins concernés : l'agrégation en 3 classes de la carte de sensibilité à la battance disponible à l'échelle nationale qui a été effectuée dans le cadre de l'application d'ARPEGES présentée ici conduit à une minimisation de la vulnérabilité intrinsèque au ruissellement sur des zones où ce risque est reconnu (typiquement Pays de Caux). L'utilisation d'une carte actualisée et à une résolution plus fine paraît nécessaire dans de telles zones pour assurer la pertinence des résultats,
- la carte du **drainage agricole** utilisée ici est basée sur le recensement général agricole de 2000, et demanderait à être actualisée,
- les cartes relatives aux **ripisylves et bandes enherbées** devraient être mises à jour au fur et à mesure de leur mise en œuvre.

7.3 Vers une agrégation des résultats pour faciliter une interprétation plus globale ?

Actuellement, les résultats sont affichés par substance active, pour une gamme de substances jugées d'intérêt pour les utilisateurs consultés (essentiellement les Agences de l'Eau). Cela permet de mieux cerner les potentiels de contamination, les causes et actions de remédiation possibles pour les substances à enjeu sur chaque bassin.

Toutefois, une vision plus globale de la contamination intéresse également les gestionnaires, notamment pour juger plus globalement de l'état général et de l'évolution de l'état des lieux. Dans ce sens, une agrégation des résultats par grandes classes de mobilité des molécules pourrait être envisagée, éventuellement croisée avec le type d'usages des molécules. Le souhait d'une agrégation plus forte, toutes substances confondues, est toujours possible mais deux limites sont à considérer : la première est qu'une telle agrégation complexifie fortement l'interprétation finale et ce que l'on peut en tirer en termes d'identification des causes et actions à mettre en œuvre pour les limiter, la seconde est qu'au final le résultat global risque de tendre à mettre en évidence un potentiel de contamination lissé et donc peu discriminant.

En conclusion, l'application d'ARPEGES par les Agences ouvre la possibilité pour celles-ci d'adapter au mieux l'outil à leurs besoins, de tester différents types d'application (adaptation des seuils de discrétisation, utilisation de données avec une meilleure résolution, agrégation des substances par famille). La gamme des possibilités semble très vaste, pour autant que les fondements de la méthode ARPEGES et de l'interprétation des résultats associés restent en tête. La mise à disposition de données actualisées au niveau national ainsi que la discrétisation de nouvelles données par les Agences sembleraient utiles, mais cela dépasse le cadre de ce travail et des missions d'Irstea.

Sigles & Abréviations

AMM : Autorisation de Mise sur le Marché

ARPEGES : Analyse de Risque Pesticides pour la Gestion des Eaux de Surface

AURELHY : Analyse Utilisant le RELief pour l'Hydrométéorologie

BDAT : Base de Données des Analyses de Terre

BDGSF : Base de Données Géographique des Sols de France

BNV-d : Base Nationale de Ventes des Distributeurs

BVME : Bassin Versant de Masse d'Eau

CLC : Corine Land Cover

Corpen : Comité d'ORientation pour des Pratiques agricoles respectueuses de l'ENvironnement

DCE : Directive Cadre sur l'Eau

EDL : Etat Des Lieux

HER : Hydro-Eco-Région

IDPR : Indice de Développement et de Persistance de Réseaux

IFT : Indice de Fréquence de Traitement

NQE : Norme de Qualité Environnementale

OTEX : Orientation Technico-Economique des Exploitations

PRA : Petite Région Agricole

PRN : Petite Région Naturelle

RGA (RA) : Recensement (Général) Agricole

RICA : Réseau d'Information Comptable Agricole

RPG : Registre Parcellaire Agricole

SAU : Surface Agricole Utile

SIRIS : Système d'Intégration du Risque par Interaction des Scores

UCS : Unité Cartographique de Sol

UTS : Unité Typologique de Sol

Bibliographie

Canellas et al, 2014. Les normales climatiques spatialisées Aurelhy 1981-2010 : températures et précipitations. Association Météo et Climat.

Carles et al, 2015. Mieux suivre spatialement l'usage des pesticides, en particulier sur les bassins versants, en s'appuyant sur un observatoire des ventes détaillé au code postal de l'utilisateur final de produit. Rapport du programme Ecophyto, 47 p

Champeaux, J.-L. and A. Tamburini, 1996. Zonage climatique de la France à partir des séries de précipitations (1971-1990) du réseau climatologique d'état. La météorologie. 8ème série **14**(Juin 1996): 45-54.

Gauroy, C., 2009. Surveillance des produits phytosanitaires dans les eaux de surface : Interprétation des données. Rapport scientifique Cemagref, 88 p.

Lagacherie, P., 1987. Carte de France de l'hydromorphie à l'échelle des petites régions naturelles, carte et notice explicative, INRA.

Le Bissonnais, Y. et al., 2002. L'érosion hydrique des sols en France. Rapport INRA, IFEN.

Mellac et Piffady, 2018. Guide d'installation et d'utilisation de l'interface informatique associée à la méthode ARPEGES. 79pp

Sanders, L. (2001). Modèles en analyse spatiale. Paris, Hermès Science. Van Genuchten, M. T., F. J. Leij and S. R. Yates, 1991. The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils, Version 1.0, EPA Report 600/2-91/065, U.S. Salinity Laboratory, USDA, ARS, Riverside, California.

Valette et al, 2008. Système Relationnel d'Audit de l'Hydromorphologie des Cours d'Eau SYRAH CE. Principes et méthodes de la sectorisation hydromorphologique.

Vernier, F., F. Zahm, F. Saudubray, K. Petit and G. Bousquet, 2009. Une méthode de spatialisation des activités agricoles au service des politiques agri-environnementales (RA-SPACE); Application au calcul d'un indicateur pesticide spatialisé sur le bassin Adour-Garonne dans le cadre de la mise en œuvre de la Directive cadre sur l'eau. Ingénieries-EAT: 39-54.

Wasson, J. G., A. Chandesris, H. Pella and L. Blanc, 2002. Les hydro-écorégions de France métropolitaine, approche régionale de la typologie des eaux courantes et éléments pour la définition des peuplements de référence d'invertébrés Cemagref. Juin 2002

8. Table des illustrations

Figure 1 : Proportion de quantification des substances actives étudiées à l'échelle nationale sur la période 2013-2015	12
Figure 2 : Pourcentage du nombre de quinzaines de jours où plus de 10 mesures ont été effectuées par HER (moyenne sur la période 2013-2015 et pour les quinze substances actives étudiées)	15
Figure 3: Illustration de la notion de bassin versant de la masse d'eau (BVME). Données : Agence de l'eau Loire-Bretagne	17
Figure 4 : Schéma très simplifié du modèle ARPEGES (les cadres gris correspondent aux données d'entrée du modèle, les cadres de couleur correspondent aux variables calculées par le modèle).....	19
Figure 5 : Schéma simplifié du modèle ARPEGES (à droite : 5a - nappe haute ; à gauche : 5b - nappe basse)	22
Figure 6 : Construction de la structure graphique d'ARPEGES - Choix des nœuds parents et enfants	28
Figure 7 : Construction de la structure graphique d'ARPEGES - Choix des états pris par les nœuds parents et enfants	29
Figure 8 : Construction de la structure graphique d'ARPEGES - Etablissement des liens entre les nœuds parents et enfants.....	30
Figure 9 : Liens les cartes de variables d'entrée et la carte de vulnérabilité intrinsèque au ruissellement (exemple nappe basse).....	36
Figure 10 : Liens les cartes de variables d'entrée et la carte de vulnérabilité intrinsèque aux écoulements de subsurface (exemple nappe basse).....	37
Figure 11 : Liens les cartes de variables d'entrée et la carte de vulnérabilité intrinsèque au drainage agricole (exemple nappe basse)	38
Figure 12 : Liens entre la carte de vulnérabilité spécifique via les transferts rapides et les cartes de ses nœuds parents (exemple glyphosate nappe basse)	41
Figure 13 : Liens entre la carte de vulnérabilité spécifique via les transferts lents et les cartes de ses nœuds parents (exemple glyphosate nappe basse)	42
Figure 14 : Extrait de la table de probabilités du noeud Vulnérabilité intrinsèque au ruissellement	44
Figure 15 : Construction de la structure probabiliste d'ARPEGES	47
Figure 16 : Concepts d'UCS et d'UTS (INRA)	50
Figure 17 : Comparaison des cartes avant (à gauche) et après (à droite) discrétisation de la donnée battance pour la zone Bretagne-Normandie	51
Figure 18 : Carte des 6 zones climatiques définies pour la méthode	54
Figure 19 : Illustration des trois cas d'agrégation possibles	60
Figure 20 : Schématisation du calcul des surfaces de drainage agricole à l'échelle du BVME par la méthode de désagrégation-agrégation.	61
Figure 21 : Schématisation de l'approche d'agrégation probabiliste.....	63
Figure 22 : Graphe du réseau ARPEGES en saison nappe basse	66
Figure 23 : Graphe du réseau ARPEGES en saison nappe haute	67
Figure 24 : Illustration sur le graphique du réseau ARPEGES du calcul pour un BVME exemple (valeurs des nœuds parents renseignées).....	68
Figure 25 : Cartes de vulnérabilité intrinsèque au ruissellement	76
Figure 26 : Carte de vulnérabilité intrinsèque aux écoulements de subsurface	77
Figure 27 : Cartes de vulnérabilité intrinsèque au drainage agricole.....	78
Figure 28 : Vulnérabilité spécifique aux transferts lents en période de Nappe Basse.....	79
Figure 29 : Illustration de la méthode de calcul de l'intégrale "moyennée" par le nombre de quinzaines	82
Figure 30 : Découpage national en HER modifiées par la prise en compte de l'occupation du sol (OTEX).....	83
Figure 31 : Position et dénomination des HER modifiées par la prise en compte de l'occupation du sol (OTEX).....	83
Figure 32 : Comparaison des cartes ARPEGES et de données observées pour le 2,4-D, le 2,4-MCPA, l'Aminotriazole et la Bentazone	85
Figure 33 : Comparaison des cartes ARPEGES et de données observées pour le Boscalid, le Chlortoluron, le Diflufénicanil et le Glyphosate.....	86
Figure 34 : Comparaison des cartes ARPEGES et de données observées pour le Métazachlore, le Nicosulfuron, la Pendiméthaline et le S-métolachlore.....	87
Figure 35 : Analyse conjointe des cartes du potentiel de contamination, de son indice de confiance, de la somme des probabilités des classes fort et très fort, et de la probabilité de la classe très fort	90

Figure 36 : Analyse des causes du potentiel de contamination « fort » sur deux exemples (Bretagne à gauche et Coteaux aquitains à droite).....	92
Figure 37 : Analyse des causes du potentiel de contamination « moyen » du S-métolachlore (légende de l'encadré en Figure 36).....	93
Figure 38 : Analyse des causes du potentiel de contamination « moyen » du Boscalid (légende de l'encadré en Figure 36).....	94
Figure 39 : Comparaison de la vulnérabilité intrinsèque des trois types d'écoulement en les deux zones "focus"	94
Figure 40 : Comparaison des cartes de potentiel de contamination via les transferts lents (à gauche) ou rapides (à droite) pour le S-métolachlore	95
Figure 41 : Analyse de la corrélation des superficies agricoles à l'échelle du canton estimées selon deux sources de données différentes : CLC 2000 et le RA 2000	117

Tableau 1 : Liste des substances actives retenues pour l'étude et classement selon les critères de sélection.....	11
Tableau 2 : Description des caractéristiques physico-chimiques des substances étudiées et de leurs usages	14
Tableau 3 : Etats des nœuds parents et seuils de discrétisation	33
Tableau 4 : Synthèse des données d'entrée d'ARPEGES.....	49
Tableau 5 : Quantités orphelines et ratios pour les substances du projet	57
Tableau 6 : Bilans sur les quantités agricoles, non-agricoles et les pertes de quantités après agrégation à l'échelle BVME	58
Tableau 7 : Coefficients de saisonnalité retenus pour le "projet ARPEGES EDL 2019"	59
Tableau 8 : Extrait de la table d'entrée pour le S-métolachlore en nappe basse.....	64
Tableau 9 : Extrait d'une table de résultat (S-métolachlore nappe basse)	65
Tableau 10 : Données d'entrée pour le BVME exemple (IDnum = 100, S-métolachlore nappe basse).....	69
Tableau 11 : Données de sortie pour le BVME exemple (IDnum = 100, S-métolachlore nappe basse)	69
Tableau 12 : Extrait du tableau 6 retravaillé	70
Tableau 13 : Répartition des quinze substances actives étudiées parmi les 9 combinaisons DT_{50}/K_{oc}	80
Tableau 14 : Analyse des deux cas de figure pour le potentiel de contamination « fort » pour le S-métolachlore	91
Tableau 15: Analyse des deux cas de figure pour le potentiel de contamination « moyen » pour le S-métolachlore et le boscalid	93

9. Annexes

9.1. Annexe 1 : Méthode d'établissement des coefficients de saisonnalité

Il s'agit dans cette annexe d'une proposition de méthode pour mobiliser l'expertise existante sur la saisonnalité des applications de produits phytosanitaires.

9.1.1. Rappel définition

Le modèle Arpèges produit des cartes de potentiel de contamination des masses d'eau de surface en distinguant deux saisons caractérisées par des types de fonctionnement hydrologique distincts (période de nappe haute et période de nappe basse, les dates limitant ces deux périodes étant variables selon les bassins en lien avec les conditions hydro-météorologiques).

La connaissance de la répartition des usages pour chaque substance entre ces deux périodes est essentielle pour calculer une pression d'usage phytosanitaire par saison, et in fine ajuster le potentiel de contamination associé à chacune des saisons.

Concrètement, il s'agit donc d'estimer quelle part de la quantité totale annuelle de la substance active est appliquée en nappe haute et en nappe basse. ARPEGES n'étant pas un modèle mécaniste, des effets de frontière de saisons peuvent biaiser l'évaluation du risque de transfert. Ainsi, des adaptations sont à prévoir pour certaines cultures faisant l'objet d'applications en frontière de saisons, afin de bien faire correspondre pression d'usage et risque de transfert.

9.1.2. Précisions sur les termes « nappe haute » et « nappe basse »

Ces qualificatifs renvoient à des fonctionnements hydrologiques pour les écoulements vers les eaux de surface, et ne font pas référence à la prise en compte de transferts d'eau vers le compartiment souterrain. Les nappes hautes ou basses auxquelles il est fait référence sont des nappes temporaires superficielles qui se forment dans certaines conditions de sols et de saturation en période hivernale. En nappe haute, le drainage par drains enterrés ou dans les fossés est actif, les sols sont saturés et la réserve utile est pleine ou quasiment pleine. Les dates bornant ces deux périodes dépendent de la zone géographique et du climat de l'année, mais ne sont pas formalisées dans le modèle, il s'agit de parler, en ce qui concerne la pression d'usage phytosanitaire, d'applications en période de saturation des sols ou d'applications en période de non-saturation.

9.1.3. Méthode de détermination des coefficients de saisonnalité

La méthode générale consiste à étudier dans quel cas se trouve la substance active :

1. si la substance active est appliquée sur une seule période, du fait de sa fonction (ex : fongicide) ou des cultures pouvant être traitées (ex : maïs)
2. si la substance active peut être appliquée sur les deux périodes, mais indépendamment des cultures traitées (traitement de sols, traitements généraux...)
3. si la période dépend de la culture traitée (auquel cas il faut inventorier les cultures pouvant être traitées par cette substance active, et déterminer les coefficients en fonction des quantités appliquées respectivement sur chaque culture)

4. [cas particulier du cas n°3] si la substance peut être appliquée sur les deux périodes pour une ou plusieurs cultures (auquel cas il faut estimer la part des quantités entre les deux périodes pour ces cultures).

Coefficients nationaux :

Etant donné le travail déjà important pour déterminer les coefficients de saisonnalité pour les 15 substances actives de l'étude, ces coefficients sont déterminés à l'échelle nationale. Cela ne signifie pas que les dates d'application sont les mêmes partout en France pour un couple SA-culture donné, mais bien que la position de l'application par rapport aux changements de fonctionnement hydrologique local des sols est la même partout en France.

Méthode particulière pour les applications en frontière de périodes :

Un cas particulier est à traiter de manière spécifique : lorsque l'application se fait en toute fin de période de « nappe basse ». Dans ce cas, dès lors que la demi-vie de la substance active est suffisamment longue, des transferts importants peuvent avoir lieu en période de « nappe haute », tout du moins pour les transferts diffus sur un temps long (« transferts lents » dans ARPEGES). C'est le cas de nombreux herbicides, qui sont souvent appliqués en pré-semis, post-semis, pré-levée ou post-levée précoce de cultures d'hiver (« entrée hiver »). C'est pourquoi on considérera que ces applications sont en « nappe haute ».

Pour l'autre cas particulier, lorsque l'application se fait en tout début de période de « nappe basse » (applications « sortie hiver »), il est possible que les transferts soient très importants lorsque des événements climatiques exceptionnels surviennent. On considère pour la présente étude que ces cas soient trop exceptionnels pour en tenir compte, aussi les applications de sortie hiver seront considérées comme étant à 100% en nappe basse.

Méthode de pondération pour les cas n°3 et 4 :

Comme indiqué plus haut pour les cas n°3 et 4 (coefficients qui dépendent des cultures), il faut pouvoir pondérer les coefficients **par l'importance des quantités appliquées** respectivement pour chaque culture et chaque saison. Or il est assez difficile pour l'instant d'avoir l'information dans ce degré de finesse à partir de la BNVD spatialisée, car cela représenterait des jeux de données beaucoup trop lourds à traiter (la donnée obtenue auprès de l'Ineris est à l'échelle de l'îlot cultural, après agrégation des quantités à cette échelle, un îlot pouvant contenir plusieurs cultures). A défaut de cela, on peut pondérer les coefficients **par les surfaces des cultures cibles**, introduisant des approximations supplémentaires mais plus faciles à calculer. On se basera pour l'instant sur les surfaces inventoriées à l'échelle nationale dans le Recensement Agricole 2010, qui distinguent les cultures d'hiver et de printemps pour certaines cultures (notamment céréales et colza).

9.1.4. Tableau de synthèse des coefficients de saisonnalité déterminé selon cette méthode

Substance active	% en « nappe haute »	% en « nappe basse »	Cas n°	Commentaires
Boscalid	0	100	1	Pas d'ambiguïté car fongicide
Nicosulfuron	0	100	1	Herbicide appliqué uniquement sur maïs
S-métolachlore	0	100	1	Herbicide appliqué uniquement sur des cultures de printemps (betterave, maïs, sorgho, tournesol, soja et quelques cultures légumières)
2,4-MCPA	0	100	1	Herbicide anti-chardon appliqué sur cultures d'hiver et de printemps, mais seulement en sortie d'hiver pour les cultures d'hiver
Bentazone	0	100	1	Herbicide anti-dicotylédones appliqué sur de nombreuses cultures de printemps, et en sortie d'hiver pour les céréales d'hiver (STEEPLE C, CERAL B)
Chlortoluron	100	0	1	Herbicide appliqué uniquement sur céréales d'hiver (blé, orge), et uniquement en entrée hiver
Oxadiazon	0	100	1	Herbicide appliqué au printemps sur cultures pérennes (vergers, vignes) et sur tournesol
Métaldéhyde	50	50	2	Molluscicide appliqué en traitement de sols, sans indication de période
Glyphosate	50	50	2	Herbicide utilisé en traitements généraux
2,4-D	50	50	2	Herbicide utilisé en traitements généraux
Aminotriazole	50	50	2	Herbicide utilisé en traitements généraux
Imidaclopride	85	15	3	Insecticide appliqué en traitement de semences (TS) pour la betterave industrielle et fourragère et les céréales (donc à la date de semis), et en post-floraison seulement pour les arbres fruitiers
Métazachlore	60	40	3	Pour le colza d'hiver, seulement des produits appliqués en entrée hiver (NOVALL, RAPSAN TDI,...), donc considérés comme étant appliqués en « nappe haute »
Pendiméthaline	55	45	3	Rattrapages printemps pour certaines cultures d'hiver (lupin notamment) et applications précoces en hiver pour vignes et vergers à forte infestation considérées comme mineures. Applications seulement en entrée hiver pour les céréales d'hiver
Diflufénicanil	80	20	4	Il peut être appliqué en entrée d'hiver (très majoritairement, avec des produits tels que TRINITY, COMPIL, CODIX,...) ou en sortie d'hiver (minoritairement, avec KALENKO par exemple, notamment, en rattrapage ou application unique de printemps).

9.1.5. Détermination des coefficients de saisonnalité pour chaque substance active

Substances actives dans les cas n°1 et 2 :

Toutes les explications sont déjà dans le tableau de synthèse précédent.

Substances actives dans le cas n°3 :

Substance active	% en « nappe haute »	% en « nappe basse »	Cas n°	Commentaires
Imidaclopride	85	15	3	Insecticide appliqué en traitement de semences (TS) pour la betterave industrielle et fourragère et les céréales (donc à la date de semis), et en post-floraison seulement pour les arbres fruitiers

Détails du calcul :

Cultures traitées par l'imidaclopride	Surfaces RA 2010 (ha)	% de la surface totale	Saison
Avoine de printemps (TS)	28477	0,36	NB
Betteraves industrielles (TS)	383588	4,79	NB
Betteraves fourragères (Plantes sarclées fourragères) (TS)	10606	0,13	NB
Blé tendre de printemps (TS)	20763	0,26	NB
Blé dur de printemps (TS)	16936	0,21	NB
Orge et escourgeon de printemps (TS)	421329	5,26	NB
Fruits à noyaux	77446	0,97	NB
Fruits à pépins	68687	0,86	NB
Avoine d'hiver (TS)	52157	0,65	NH
Blé tendre d'hiver (TS)	4876132	60,86	NH
Blé dur d'hiver (TS)	488906	6,10	NH
Orge et escourgeon d'hiver (TS)	1153291	14,40	NH
Seigle et méteil (TS)	29842	0,37	NH
Triticale (TS)	383257	4,78	NH

Substance active	% en « nappe haute »	% en « nappe basse »	Cas n°	Commentaires
Métazachlore	60	40	3	Pour le colza d'hiver, seulement des produits appliqués en entrée hiver (NOVALL, RAPSAN TDI,...), donc considérés comme étant appliqués en « nappe haute »

Détails du calcul :

Cultures traitées par le métazachlore	Surfaces RA 2010 (ha)	% de la surface totale	Saison
Légumes frais	207642	8,64	NB
Tournesol	691870	28,80	NB
Autres plantes aromatiques, médicinales et à parfum	38942	1,62	NB
Colza de printemps (et navette)	5736	0,24	NB
Colza d'hiver (et navette)	1458133	60,70	NH

Substance active	% en « nappe haute »	% en « nappe basse »	Cas n°	Commentaires
Pendiméthaline	55	45	3	Rattrapages printemps pour certaines cultures d'hiver (lupin notamment) et applications précoces en hiver pour vignes et vergers à forte infestation considérées comme mineures. Applications seulement en entrée hiver pour les céréales d'hiver

Cultures traitées par la pendiméthaline	Surfaces RA 2010 (ha)	% de la surface totale	saison
Blé dur de printemps	16936	0,1	NB
Blé tendre de printemps	20763	0,2	NB
Légumes frais	207642	1,7	NB
Lentilles (y compris semences)	15324	0,1	NB
Maïs fourrage et ensilage (plante entière)	1387081	11,3	NB
Maïs grain	1616087	13,1	NB
Orge et escourgeon de printemps	421329	3,4	NB
Fèverole de printemps	75734	0,6	NB
Fruits à pépins	68687	0,6	NB
Vignes	788637	6,4	NB
Pois protéagineux de printemps	119377	1,0	NB
Tournesol	691870	5,6	NB
Soja	49677	0,4	NB
Sorgho	46459	0,4	NB
Féverole d'hiver	75734	0,6	NH
Avoine d'hiver	52157	0,4	NH
Blé dur d'hiver	488906	4,0	NH
Blé tendre d'hiver	4876132	39,6	NH
Orge et escourgeon d'hiver	1153291	9,4	NH
Pois protéagineux d'hiver	119377	1,0	NH
Seigle et méteil	29842	0,2	NH
Colza	pas compté, car uniquement 1 produit (ATIC AQUA) pour lutter uniquement contre le coquelicot		

Substances actives dans le cas n°4 :

Substance active	% en « nappe haute »	% en « nappe basse »	Cas n°	Commentaires
Diflufénicanil	80	20	4	

Détails du calcul :

Le diflufénicanil, appliqué sur céréales seulement (sauf orge de printemps), est donc appliqué quasi-uniquement sur des cultures d'hiver :

Cultures traitées par le diflufénicanil	Surfaces RA 2010 (ha)	% de la surface totale
Blé dur de printemps	16936	0,24
Blé tendre de printemps	20763	0,30
Blé dur d'hiver	488906	7,02
Blé tendre d'hiver	4876132	69,97
Orge et escourgeon d'hiver	1153291	16,55
Seigle et méteil	29842	0,43
Triticale	383257	5,50

Il peut être appliqué en entrée d'hiver (très majoritairement, avec des produits tels que TRINITY, COMPIL, CODIX,...) ou en sortie d'hiver (minoritairement, avec KALENKO par exemple, notamment, en rattrapage ou application unique de printemps). C'est pourquoi une répartition de 80% en « nappe haute » et 20% en « nappe basse » a été choisie.

9.2. Annexe 2 : Fiche synthétique pour estimer la variable « Ruissellement/Infiltration »

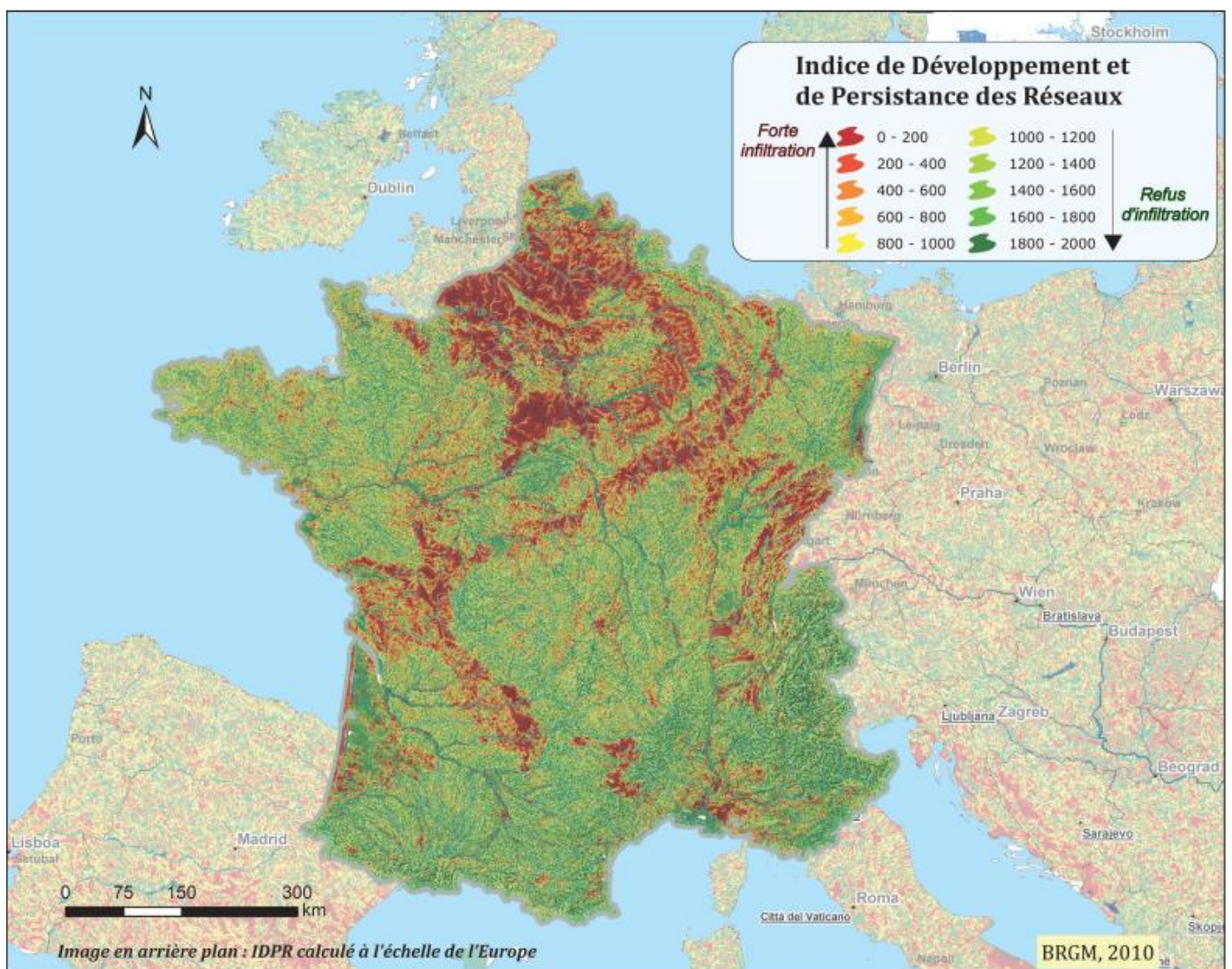
9.2.1. Objectifs

Caractériser la destination préférentielle des flux sur un BVME (eaux de surface / eaux souterraines).

9.2.2. Données sources

L'Indice de développement et de persistance de réseaux (IDPR), du BRGM (Mardhel 2006), calculé sur un maillage de 1km². L'IDPR résulte d'une comparaison entre le réseau hydrographique théorique (calculé à partir du MNT) et le réseau réel. Si le réseau réel est moins dense que le réseau théorique, cela traduit le fait que l'infiltration de l'eau est majoritaire, tandis que s'il est plus dense les flux se font majoritairement à destination des eaux de surface.

La carte de l'IDPR avant discrétisation est la suivante :



Donnée d'origine IDPR par mailles (de 50 à 5000m de largeur)

9.2.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 1 : le BVME contient les entités de la donnée

Méthode :

Agrégation : moyenne IDPR au sein d'un BVME

9.2.4. Rappel des modalités que peut prendre la variable dans ARPEGES

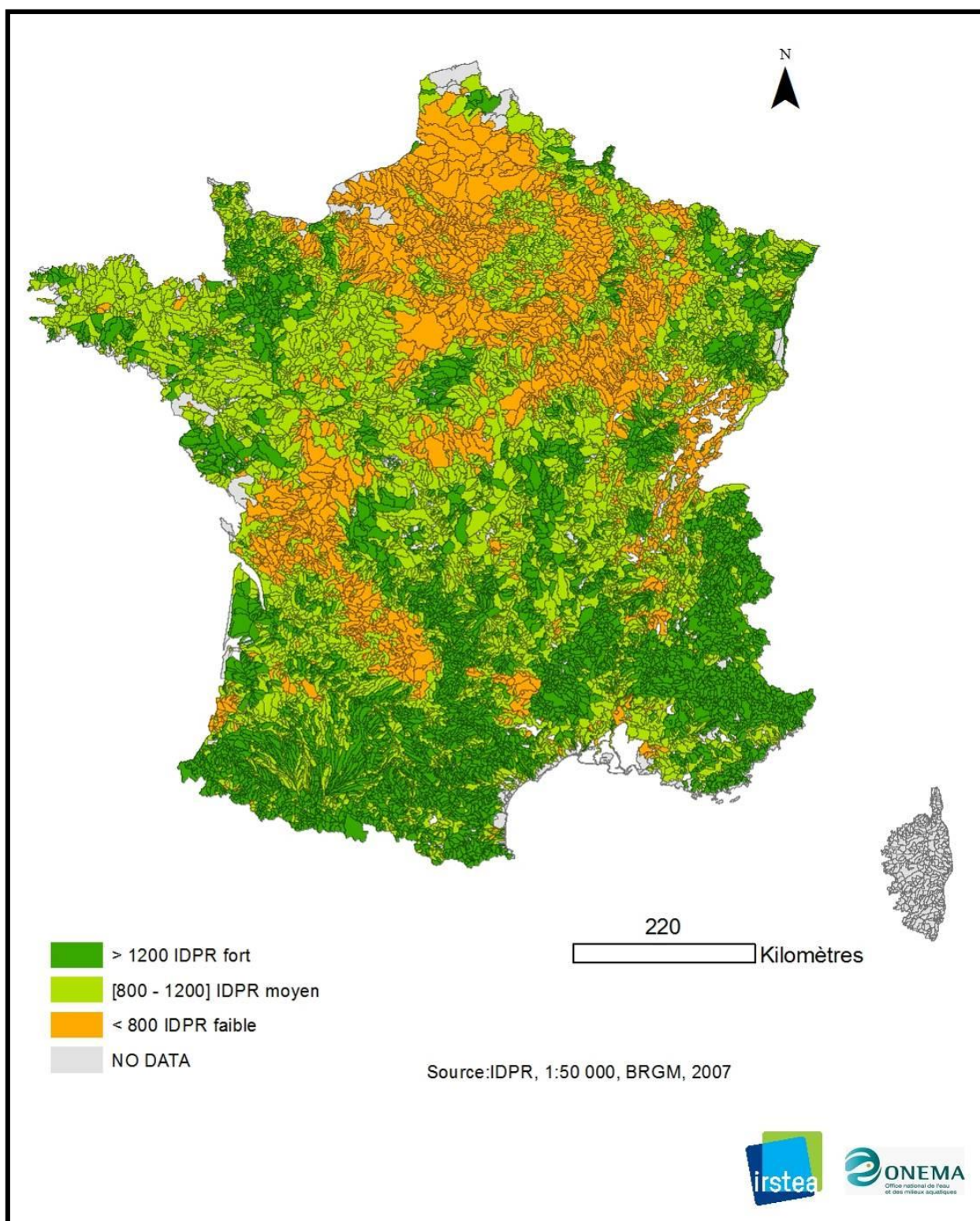
<800 → faible

Entre 800 et 1200 → moyen

>1200 → fort

9.2.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carte ci-dessous présente l'IDPR moyen sur chaque BVME avec une discrétisation selon les modalités définies dans le modèle ARPEGES. Un IDPR fort signifie une destination préférentielle vers les eaux de surface, un IDPR faible vers les eaux souterraines et un IDPR moyen une destination aussi bien vers les eaux de surface que vers les eaux souterraines.



IDPR à l'échelle du bassin versant local des masses d'eaux, discrétisé selon les modalités d'entrée du modèle ARPEGES

9.3. Annexe 3 : Fiche synthétique pour estimer la variable « Réserve utile »

9.3.1. Objectifs

Caractériser la réserve utile des sols pour chaque BVME.

9.3.2. Données sources

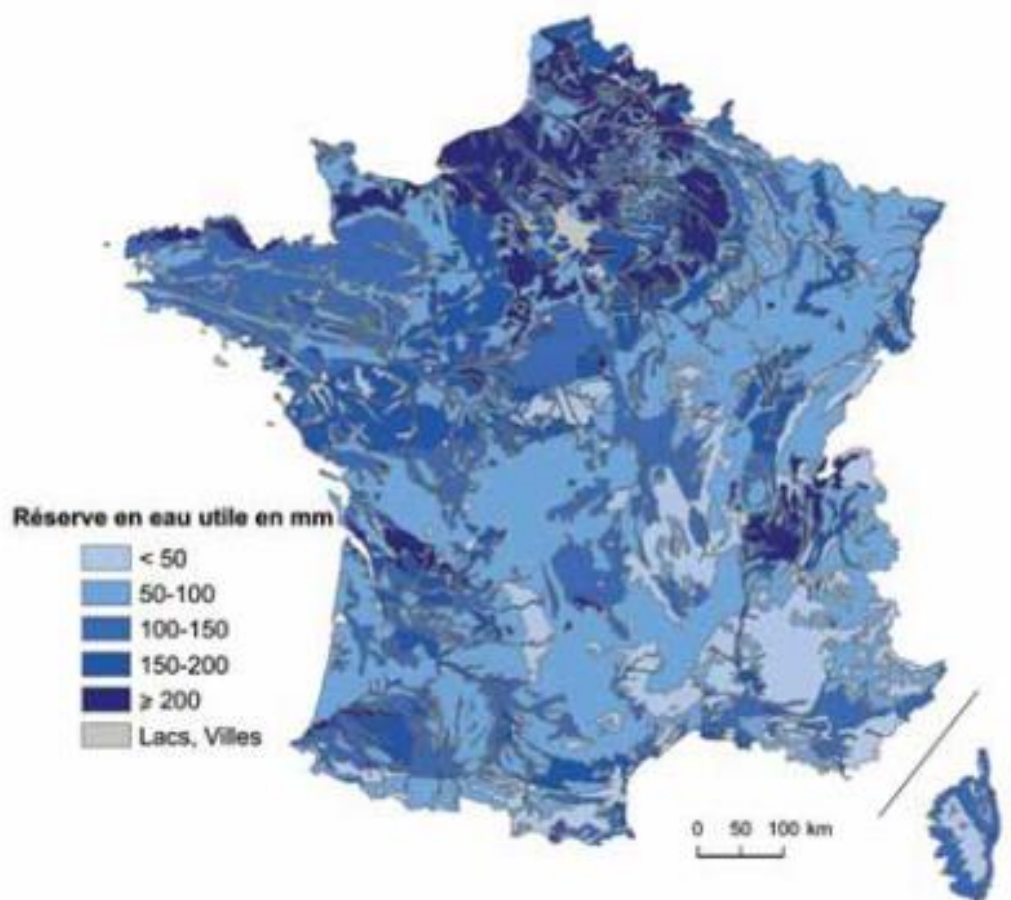
Base de données géographique des sols de France (BDGSF) au 1/1.000.000ème de l'INRA GIS-SOL.

La BDGSF est composée tout d'abord d'une carte des Unités Cartographiques de Sols (UCS). Ces unités ne sont pas homogènes, mais sont constituées d'Unités Typologiques de Sols (UTS) qui ne sont pas cartographiables à cette échelle. A chaque UCS correspond donc une liste d'UTS avec la proportion de l'UCS qu'elle représente. Chaque UTS est décrite par une quarantaine de variables comme la réserve utile (Le Bissonnais et al. 2002).

Cette réserve est exprimée en mm et discrétisée en pour chaque UTS en cinq catégories :

- < 50 mm
- 50 - 100 mm
- 100 - 150 mm
- 150 -200 mm
- > 200 mm

La carte de réserve utile avant discrétisation est la suivante :



Source : Inra, Base de données Géographique des Sols de France à 1/1 000 000, 1998.

9.3.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 2 : le BVME et les entités de la donnée se recouvrent

Méthode :

Agrégation probabiliste.

9.3.4. Rappel des modalités que peut prendre la variable dans ARPEGES

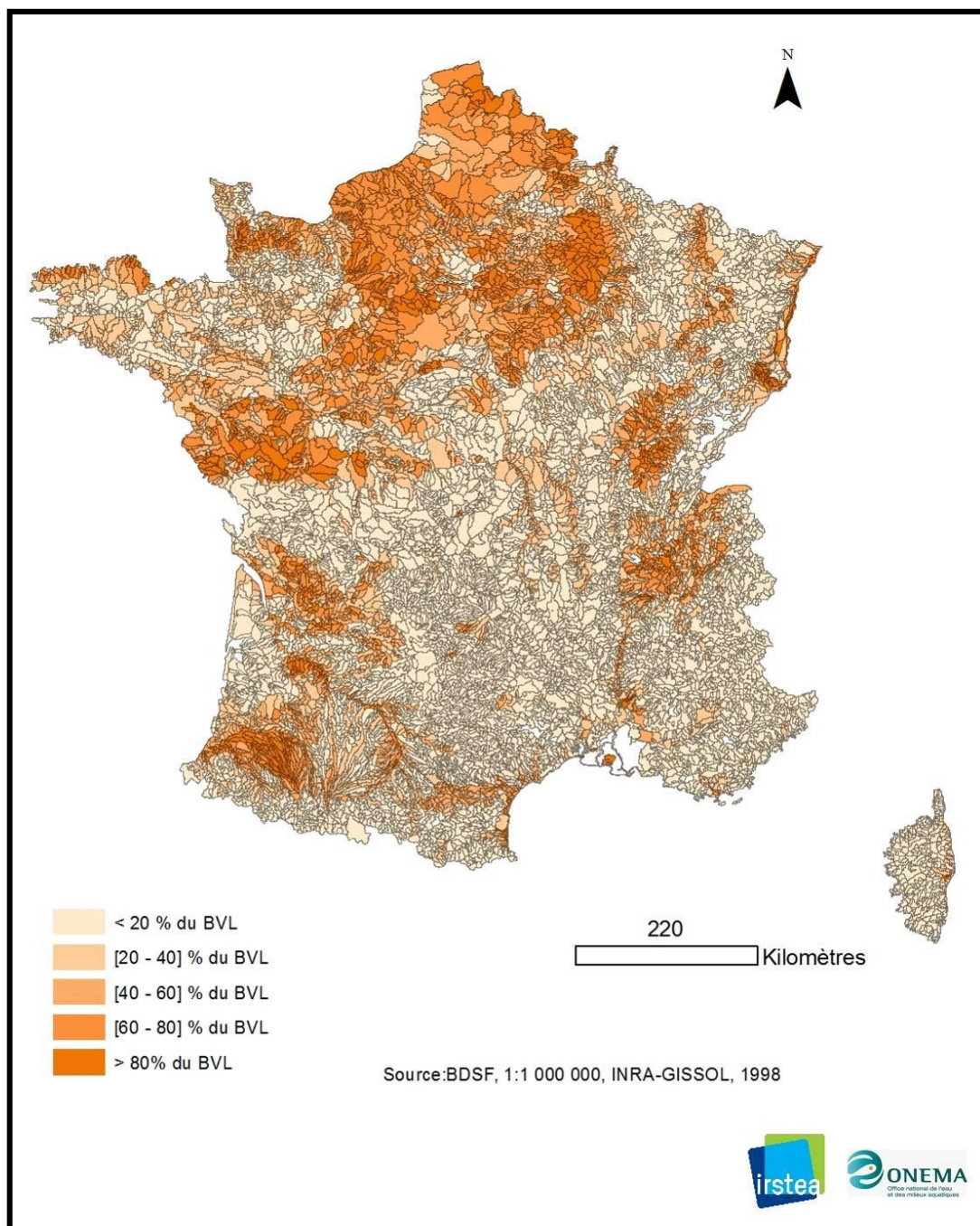
Faible qui correspond au pourcentage surfacique des catégories « < 50 mm et 50 - 100 mm » recouvrant le BVME

Moyen qui correspond au pourcentage surfacique de la catégorie « 100 - 150 mm » recouvrant le BVME

Fort qui correspond au pourcentage surfacique des catégories « 150 -200 mm et >= 200 mm » recouvrant le BVME

9.3.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

Une carte pour chaque modalité résulte de cette méthode. La carte ci-dessous présente le résultat pour la modalité forte discrétisée en 5 classes (intervalles égaux).



Présence d'une réserve utile forte à l'échelle du bassin versant local des masses d'eaux.

9.4. Annexe 4 : Fiche synthétique pour estimer la variable « drainage »

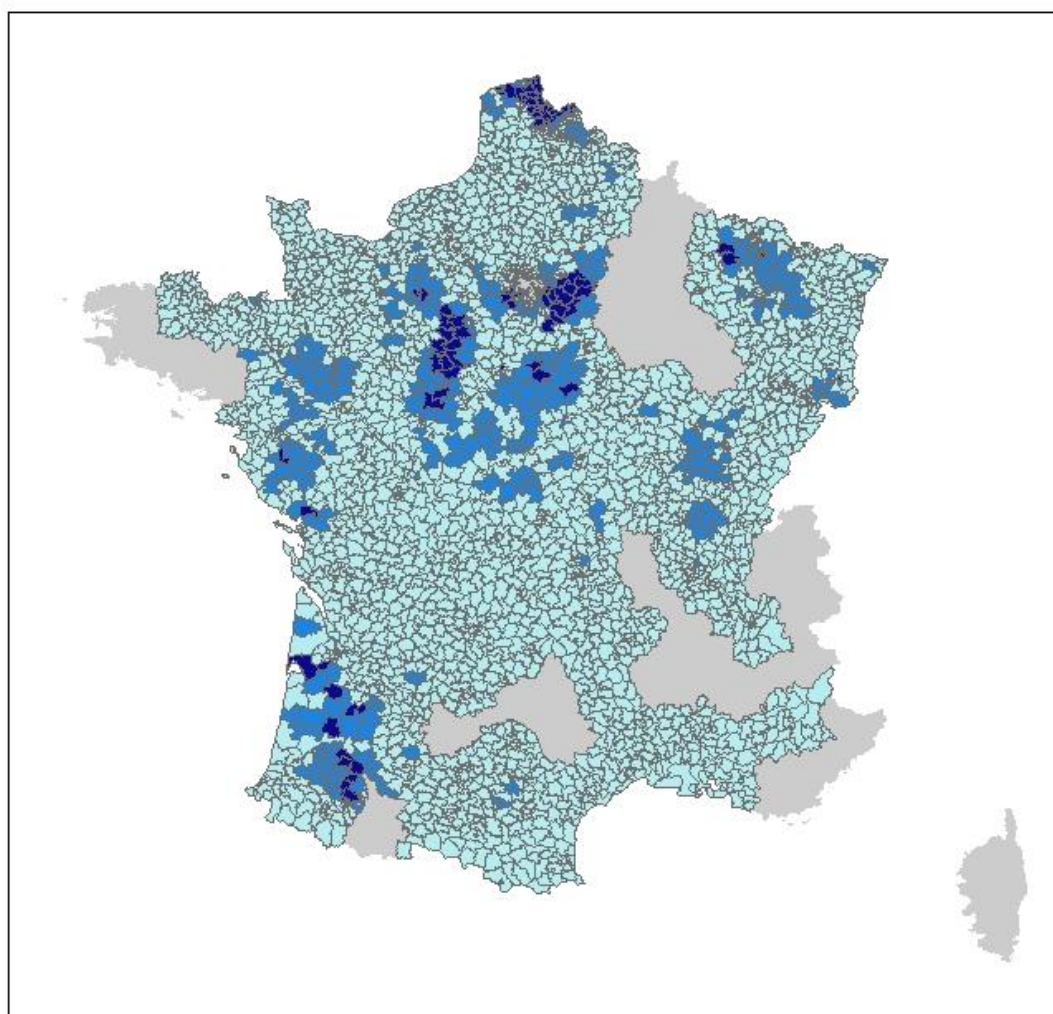
9.4.1. Objectifs

Caractériser le drainage (les superficies drainées) pour chaque BVME.

9.4.2. Données sources

RA 2000 échelle du canton

La carte de drainage avant discrétisation est la suivante :



Légende

Pourcentage de SAU drainée

	Moins de 20 %
	Entre 20 et 50 %
	Plus de 50 %
	Pas d'information

9.4.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 2 : le BVME et les entités de la donnée se recouvrent

Méthode :

Désagrégation-réagrégation

Variable de contrôle :

Superficies agricoles estimées à partir de la donnée CORINE Land Cover (CLC) 2000

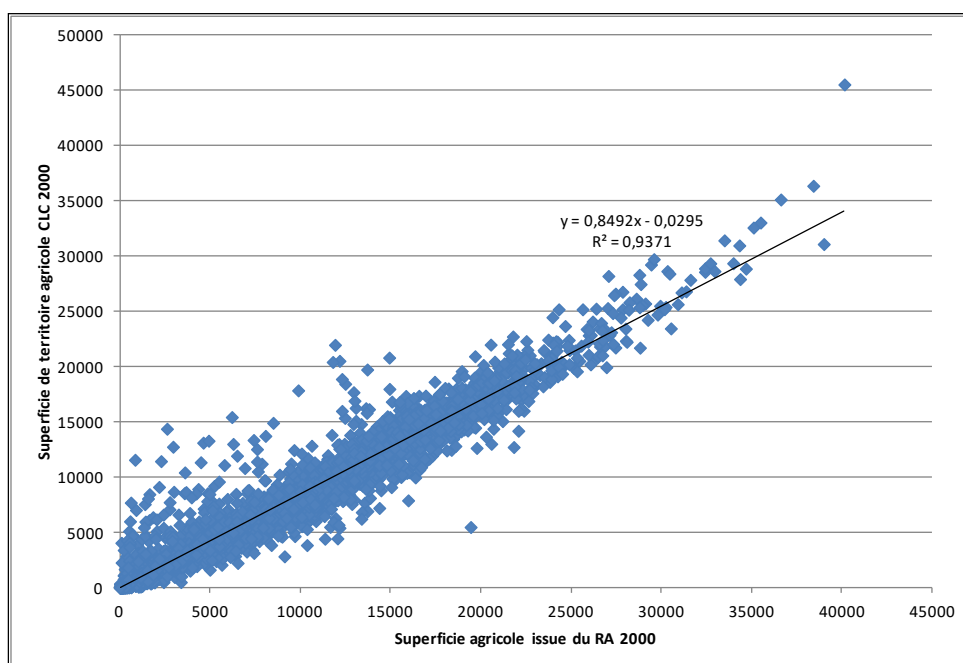
Hypothèse :

On suppose que les superficies drainées au sein d'un canton se répartissent de manière uniforme quel que soit le type de cultures et de pratiques.

Détails :

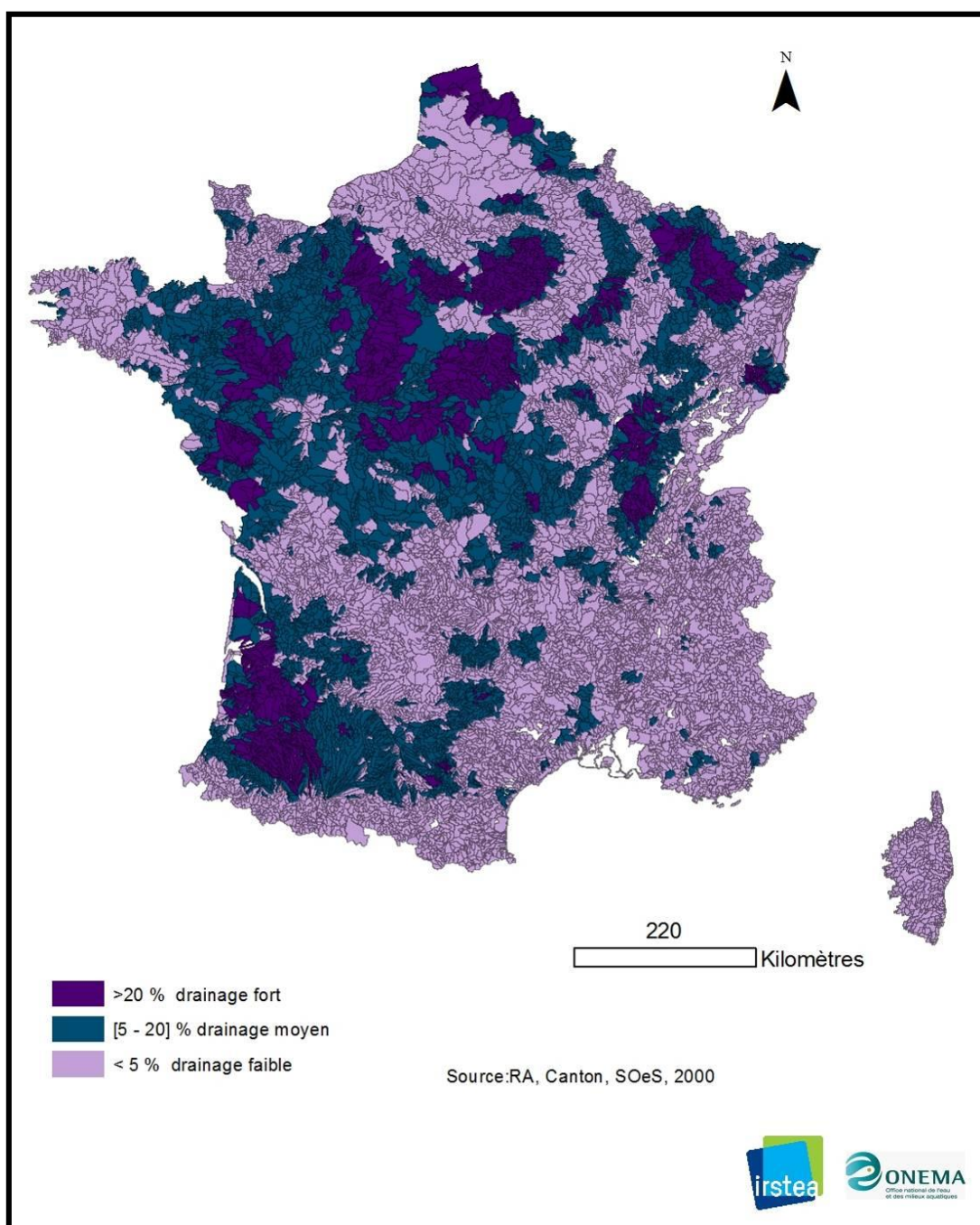
De manière à évaluer la capacité de CLC à estimer les superficies agricoles (code Corine 2), la corrélation entre la valeur de la statistique de la SAU (Surface Agricole Utile) issue du RA 2000 et celle de la superficie de territoire agricole de CLC 2000 a été analysée à l'échelle du canton. D'après les résultats présentés Figure 3.1, on observe une très bonne corrélation ($R^2=0.93$) entre la superficie de territoire agricole de CLC 2000 et celle recensée dans le RA 2000 à l'échelle du canton. CLC est donc une source de données fiable pour estimer la superficie agricole au sein d'un canton. CLC est disponible à une échelle plus précise que le RA 2000, c'est donc cette donnée qui a été utilisée.

Figure 41 : Analyse de la corrélation des superficies agricoles à l'échelle du canton estimées selon deux sources de données différentes : CLC 2000 et le RA 2000



9.4.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carte ci-dessous présente les superficies drainées par BVME discrétisées selon les modalités d'entrée du modèle ARPEGES.



Superficies drainées à l'échelle du bassin versant local des masses d'eaux discrétisées selon les modalités d'entrée du modèle ARPEGES

9.5. Annexe 5 : Fiche synthétique pour estimer la variable « Hydromorphie »

9.5.1. Objectifs

Caractériser l'hydromorphie des sols pour chaque BVME.

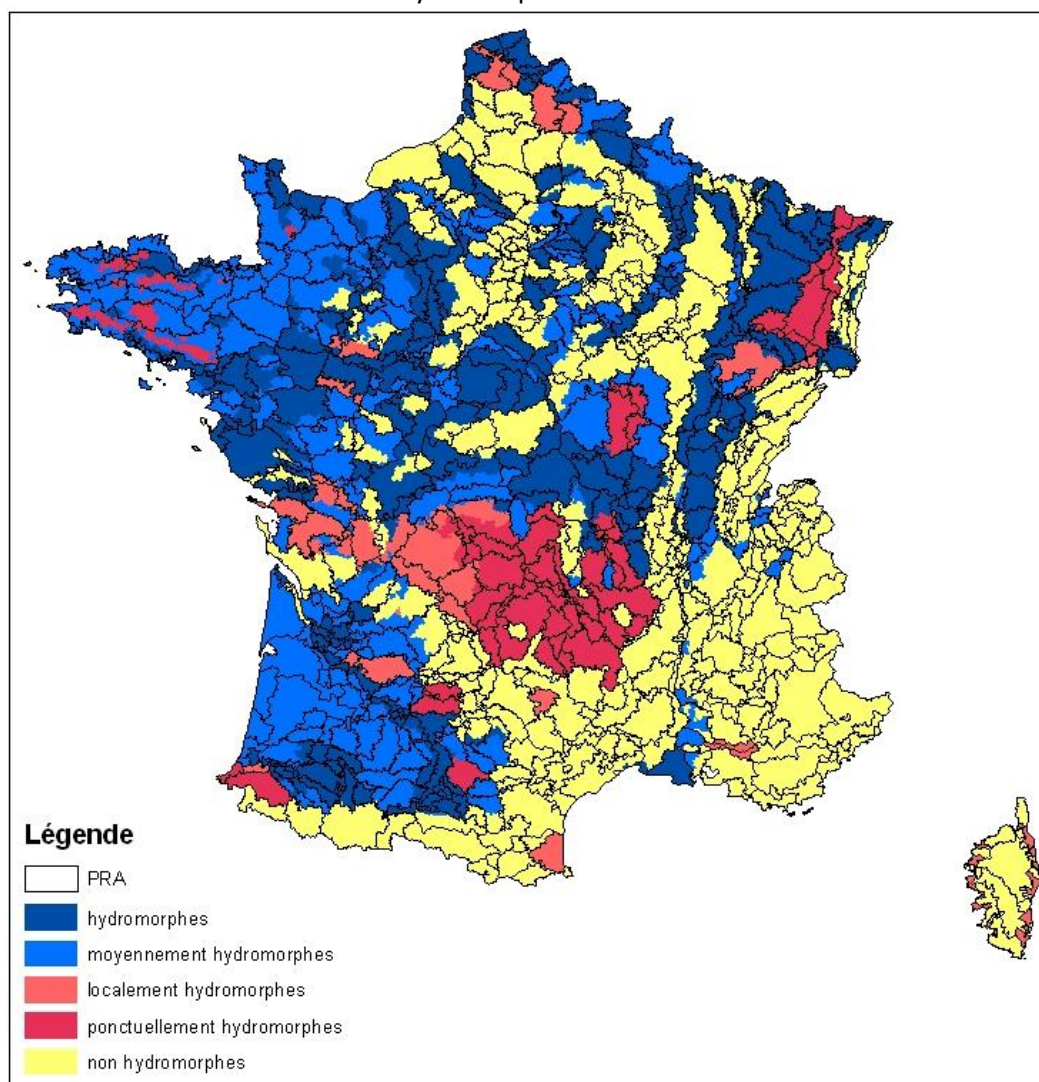
9.5.2. Données sources

Carte de France de l'hydromorphie à l'échelle des petites régions naturelles de P. Lagacherie, 1987 à l'échelle des Petites régions naturelles, qui correspondent aux petites régions agricoles éventuellement redécoupées si elles sont considérées comme trop hétérogènes.

L'hydromorphie est exprimée en pourcentage de sols hydromorphes dans la petite région naturelle discrétisé en cinq catégories :

- moins de 20 %
- moins de 20 % hydromorphe (engorgement ponctuel)
- moins de 20 % hydromorphe (+ de 5000 ha)
- 20-50 % hydromorphe
- 50 % hydromorphe

La carte avant discrétisation de l'hydromorphie est la suivante :



9.5.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 2 : le BVME et les entités de la donnée se recouvrent

Méthode :

Agrégation probabiliste.

9.5.4. Rappel des modalités que peut prendre la variable dans ARPEGES

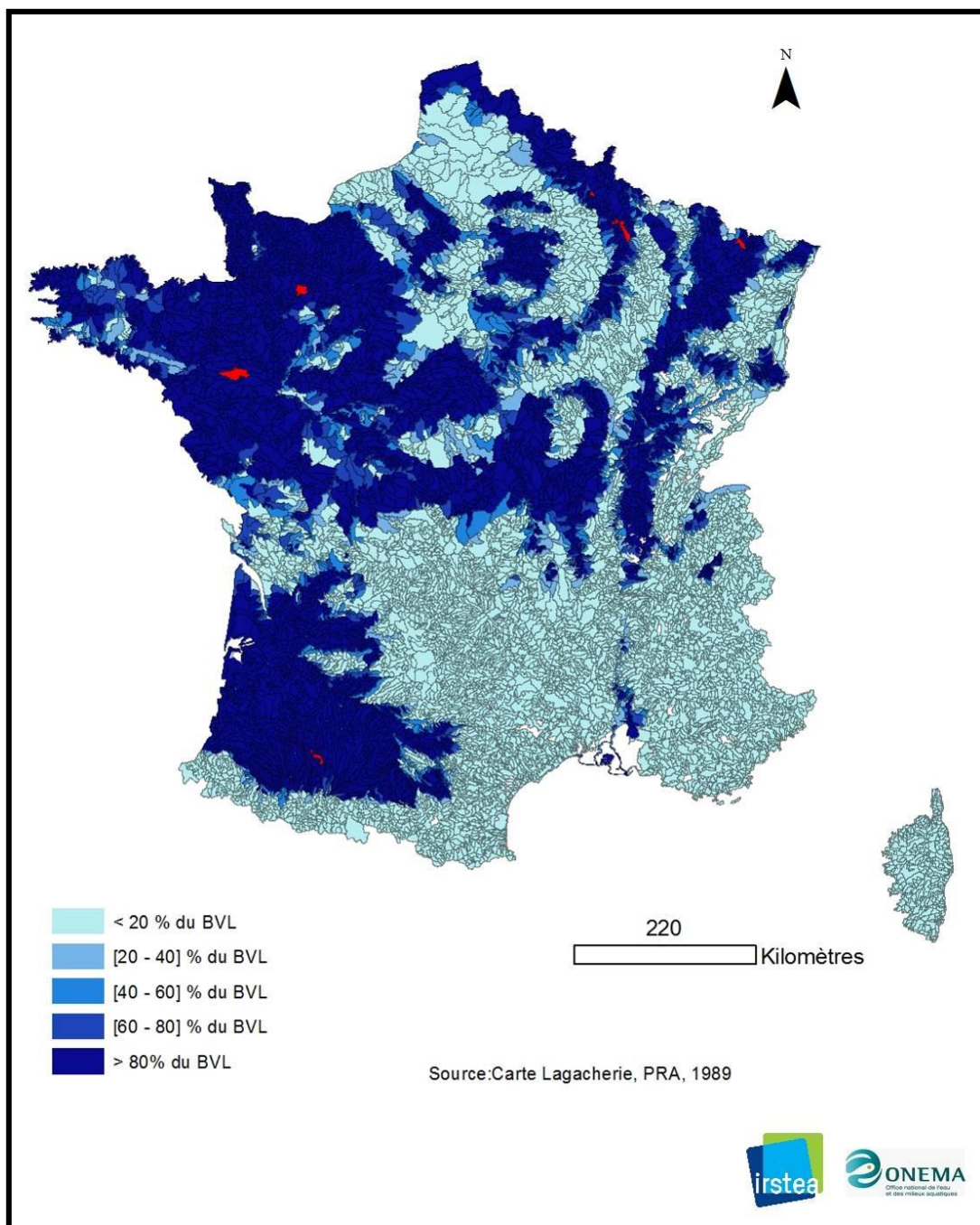
Faible qui correspond au pourcentage surfacique de la catégorie « moins de 20 % » recouvrant le BVME

Moyen qui correspond au pourcentage surfacique des catégories « moins de 20 % hydromorphe (engorgement ponctuel) + moins de 20 % hydromorphe (+ de 5000 ha) » recouvrant le BVME

Forte qui correspond au pourcentage surfacique des catégories « 20-50 % hydromorphe + 50 % hydromorphe » recouvrant le BVME

9.5.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

Une carte pour chaque modalité résulte de cette méthode. La carte ci-dessous présente le résultat pour la modalité forte discrétisée en 5 classes (intervalles égaux).



Présence d'une hydromorphie forte à l'échelle du bassin versant local des masses d'eaux.

9.6. Annexe 6 : Fiche synthétique pour estimer la variable « Battance »

9.6.1. Objectifs

Caractériser la battance des sols pour chaque BVME.

9.6.2. Données sources

Base de données géographique des sols de France (BDGSF) au 1/1.000.000ème de l'INRA GIS-SOL.

La BDGSF est composée tout d'abord d'une carte des Unités Cartographiques de Sols (UCS). Ces unités ne sont pas homogènes, mais sont constituées d'Unités Typologiques de Sols (UTS) qui ne sont pas cartographiables à cette échelle. A chaque UCS correspond donc une liste d'UTS avec la proportion de l'UCS qu'elle représente. Chaque UTS est décrite par une quarantaine de variables comme la sensibilité à la formation d'une croûte de battance obtenue par des règles de pédotransfert, (texture et du type de sol) (Le Bissonnais et al. 2002).

Cette sensibilité à la formation d'une croûte de battance est discrétisée pour chaque UTS en cinq catégories :

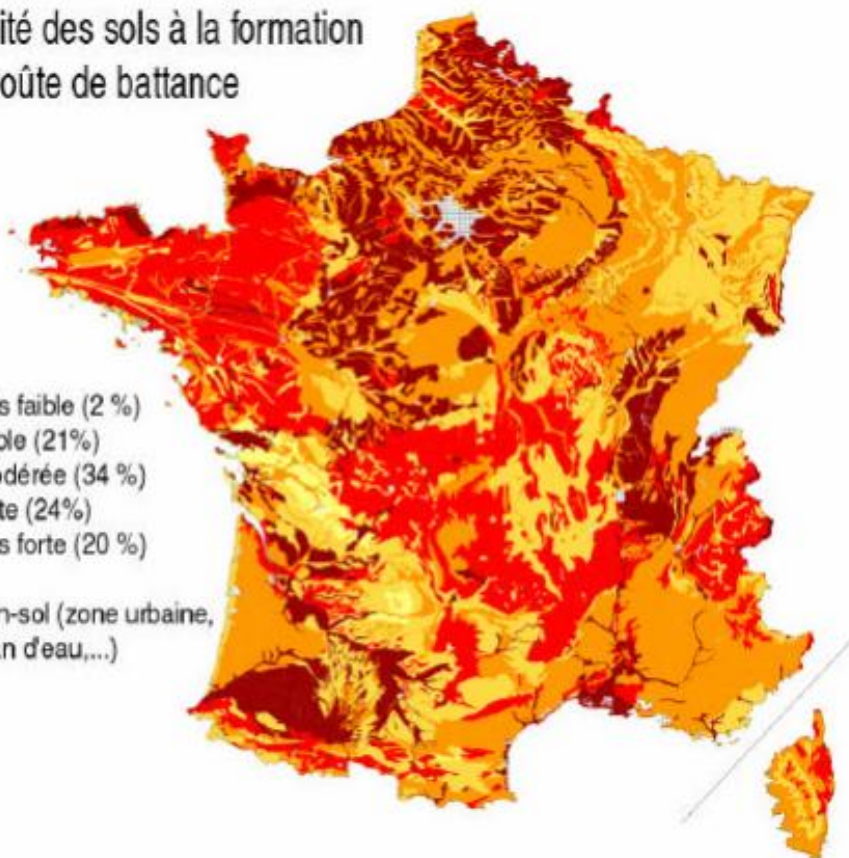
- très faible
- faible
- modérée
- fort
- très fort

La carte avant discrétisation est la suivante :

Sensibilité des sols à la formation d'une croûte de battance

- très faible (2 %)
- faible (21%)
- modérée (34 %)
- forte (24%)
- très forte (20 %)

non-sol (zone urbaine,
plan d'eau,...)



9.6.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 2 : le BVME et les entités de la donnée se recouvrent

Méthode :

Agrégation probabiliste.

9.6.4. Rappel des modalités que peut prendre la variable dans ARPEGES

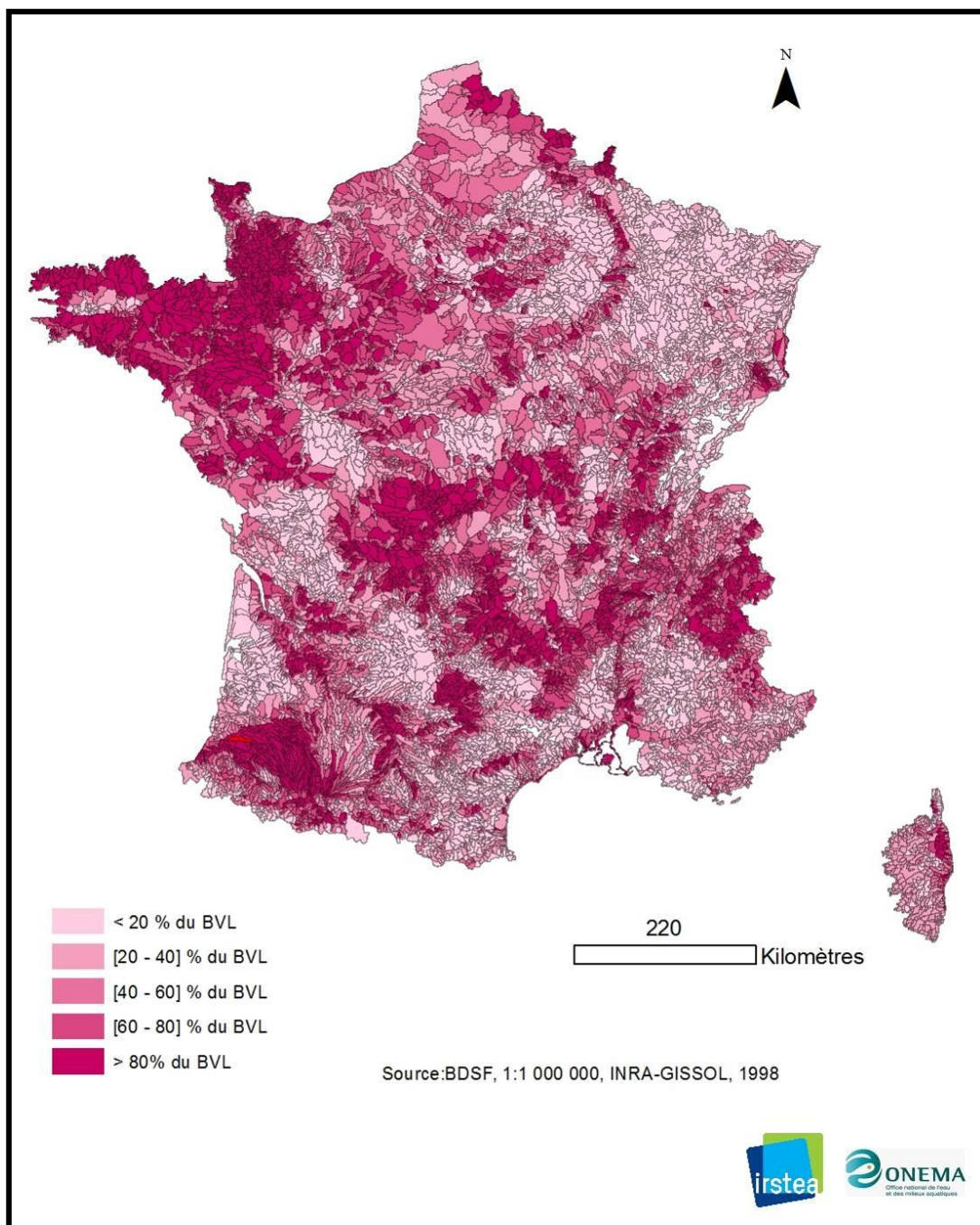
Faible qui correspond au pourcentage surfacique des catégories « faible et très faible » recouvrant le BVME

Moyen qui correspond au pourcentage surfacique de la catégorie « modérée » recouvrant le BVME

Fort qui correspond au pourcentage surfacique des catégories « forte et très forte » recouvrant le BVME

9.6.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

Une carte pour chaque modalité résulte de cette méthode. La carte ci-dessous présente le résultat pour la modalité forte discrétisée en 5 classes (intervalles égaux).



Présence d'une battance forte à l'échelle du bassin versant local des masses d'eaux

9.7. Annexe 7 : Fiche synthétique pour estimer la variable « Ripisylves 30 m »

9.7.1. Objectifs

Caractériser la présence d'une strate arborée dans la bande de 30 m (zone rivulaire effective) de part et d'autre du cours d'eau pour chaque BVME.

9.7.2. Données sources

La couche végétation de la BDTOPO® IGN de résolution métrique 2010.

9.7.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 1 : le BVME contient les entités de la donnée

Méthode :

Agrégation : pourcentage surfacique des polygones « forêt » de la couche de végétation BDTOPO découpés dans un buffer de 30 m autour du réseau hydrographique du BVME (réseau BDTOPO issu du réseau SYRAH).

9.7.4. Rappel des modalités que peut prendre la variable dans ARPEGES

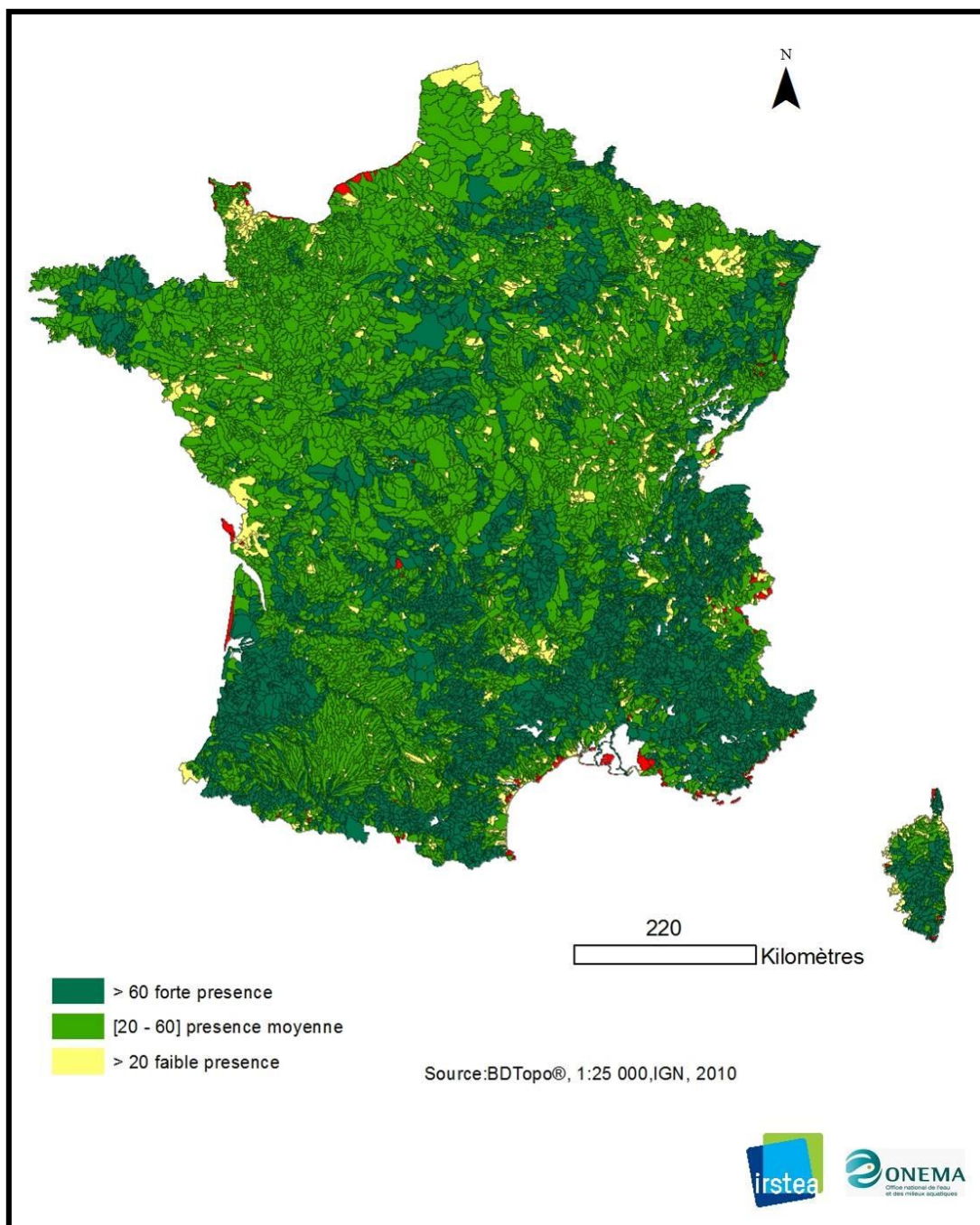
Moins de 20 % de la surface est de la forêt → faible

De 20 à 60 % → moyen

>60 % → fort

9.7.5. Résultats et limites

La carte ci-dessous illustre cette présence de ripisylves dans la bande de 30 m discrétisée selon les modalités d'entrée du modèle ARPEGES.



Présence de ripisylves dans la bande des 30 m de part et d'autre du cours d'eau à l'échelle du bassin versant local des masses d'eaux discrétisée selon les modalités d'entrée du modèle ARPEGES

9.8. Annexe 8 : Fiche synthétique pour estimer la variable « Ripisylves 10 m »

9.8.1. Objectifs

Caractériser la présence d'une strate arborée dans la bande de 10 m (haie d'arbre) de part et d'autre du cours d'eau pour chaque BVME.

9.8.2. Données sources

La couche végétation de la BDTPOPO® IGN de résolution métrique 2010.

9.8.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 1 : le BVME contient les entités de la donnée

Méthode :

Agrégation : pourcentage surfacique des polygones « forêt » de la couche de végétation BDTPOPO découpés dans un buffer de 10 m autour du réseau hydrographique du BVME (réseau BDTPOPO issu du réseau SYRAH).

9.8.4. Rappel des modalités que peut prendre la variable dans ARPEGES

Moins de 20 % de la surface est de la forêt → faible

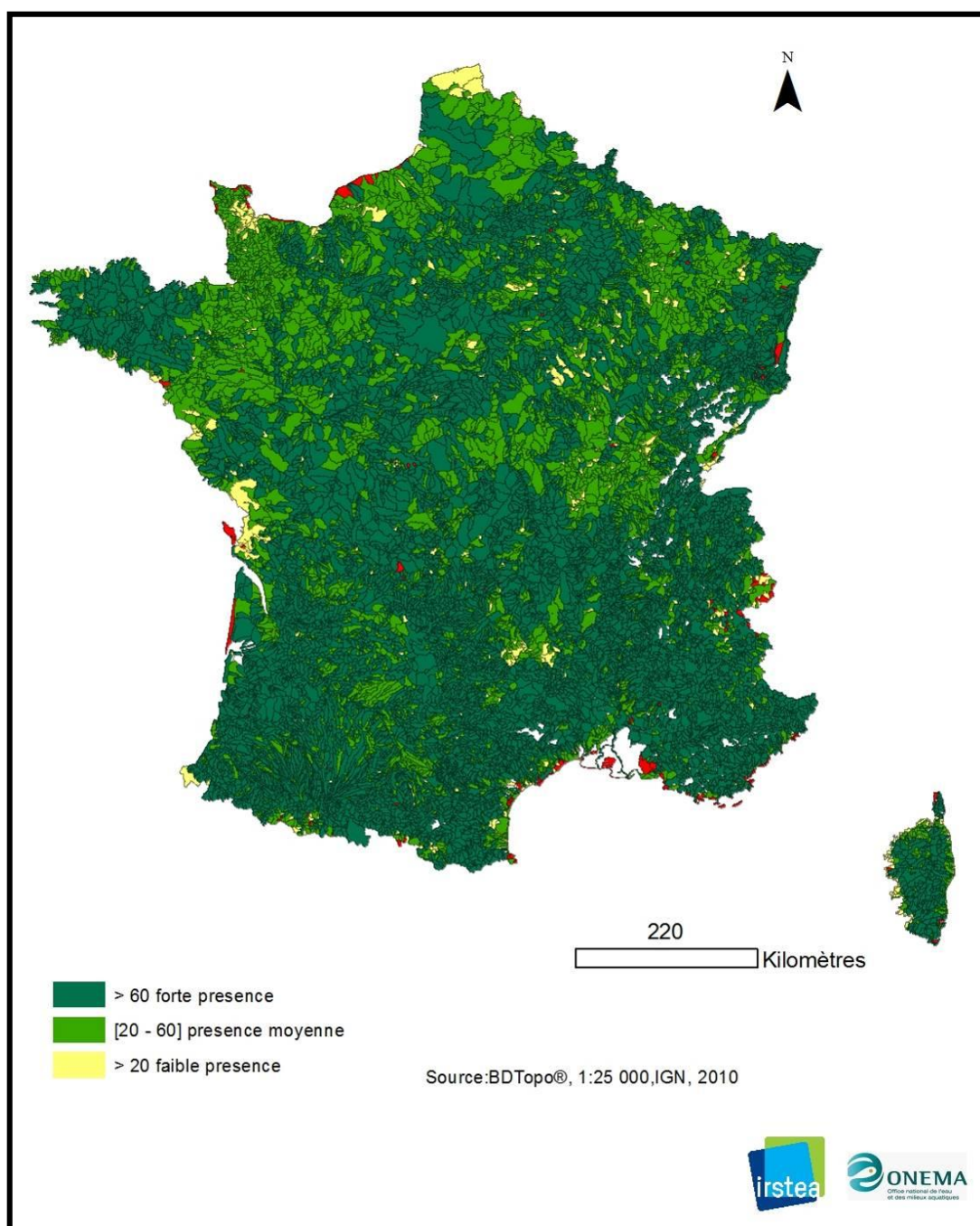
De 20 à 60 % → moyen

>60 % → fort

9.8.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carte ci-dessous illustre cette présence de ripisylves dans la bande de 10 m discrétisée selon les modalités d'entrée du modèle ARPEGES.

Cette méthode ne permet pas d'évaluer si la forêt est bien répartie le long du cours d'eau ou pas, ni si elle est située entre les zones de culture et le cours d'eau, ce qui a des conséquences différentes sur les transferts de pesticides.



Présence de ripisylves dans la bande des 10 m de part et d'autre du cours d'eau à l'échelle du bassin versant local des masses d'eaux discrétisée selon les modalités d'entrée du modèle ARPEGES

9.9. Annexe 9 : Fiche synthétique pour estimer la variable « bande enherbée »

9.9.1. Objectifs

Caractériser la présence de bande enherbée dans la bande légiférée (bande de 5 m de part et d'autre du cours d'eau) pour chaque BVME.

9.9.2. Données sources

BDTOPO® IGN de résolution métrique 2010.

Faute de données sur la présence réelle de bandes enherbées en France, nous avons considéré le pourcentage de linéaire de cours d'eau qui doit légalement être protégé par des bandes enherbées, c'est-à-dire les cours d'eau qui portent un nom sur les cartes IGN au 1/25.000ème.

9.9.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 1 : le BVME contient les entités de la donnée

Méthode :

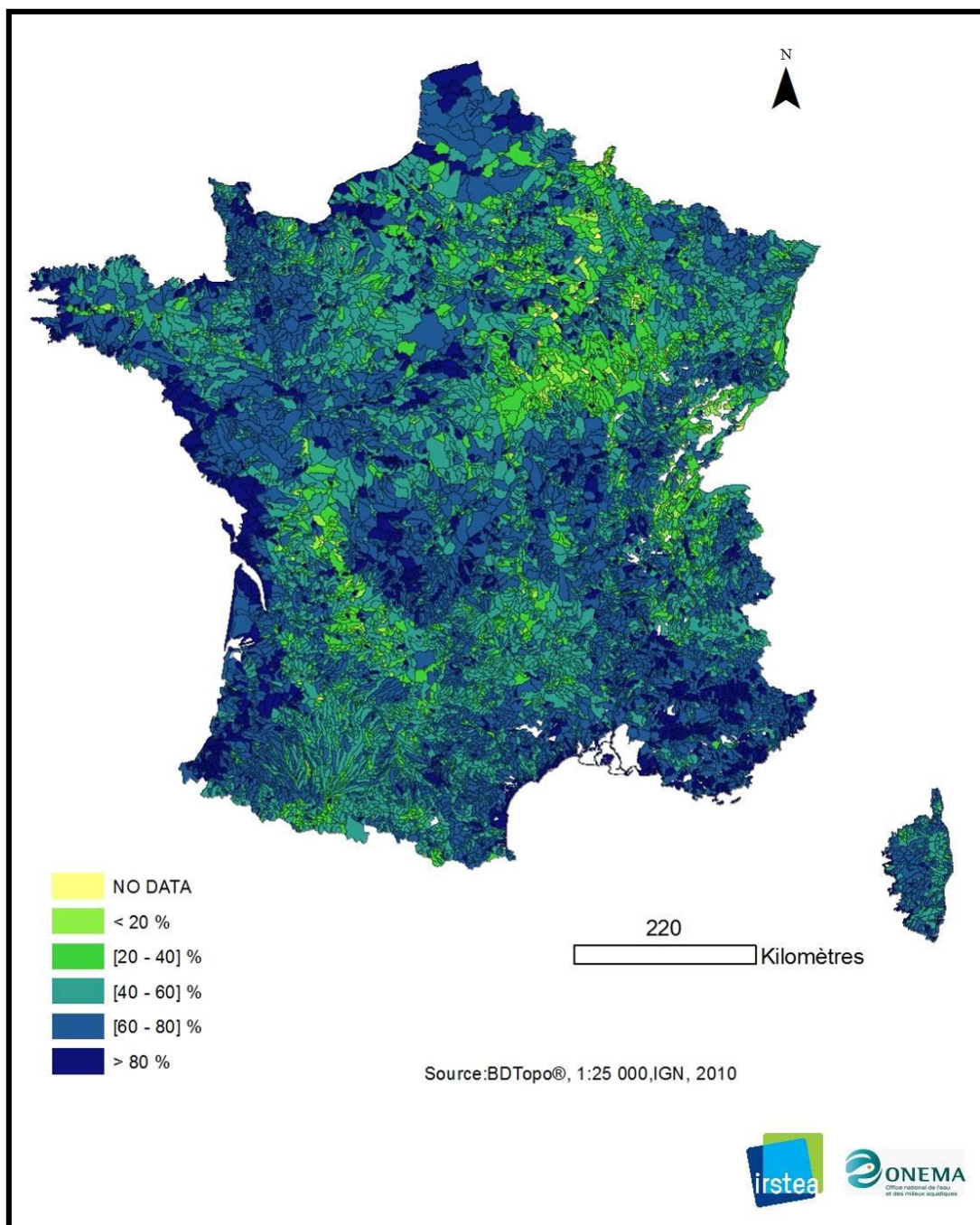
Agrégation : pourcentage linéaire de tronçon sans nom

9.9.4. Rappel des modalités que peut prendre la variable dans ARPEGES

présence qui correspond au pourcentage de cours d'eau avec un nom
absence qui correspond au pourcentage de cours d'eau sans nom

9.9.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carteci-dessous présente le résultat pour la modalité absence.



Absence de bande enherbée à l'échelle du bassin versant local des masses d'eaux

9.10. Annexe 10 : Fiche synthétique pour estimer la variable «Zone climatique»

9.10.1. Objectifs

Caractériser les paramètres climatiques qui entrent en jeu dans les phénomènes de ruissellement, notamment l'intensité, la fréquence et le cumul des pluies.

9.10.2. Données sources

Les 6 zones climatiques ont été définies à l'aide d'un zonage existant basé sur les séries de précipitations (Champeaux & Tamburini 1996). Les zones ainsi obtenues à dire d'experts correspondent à des zones globalement homogènes du point de vue de la fréquence et de l'intensité des pluies.

9.10.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 3 : le BVME est complètement contenu dans les entités de la donnée

Méthode :

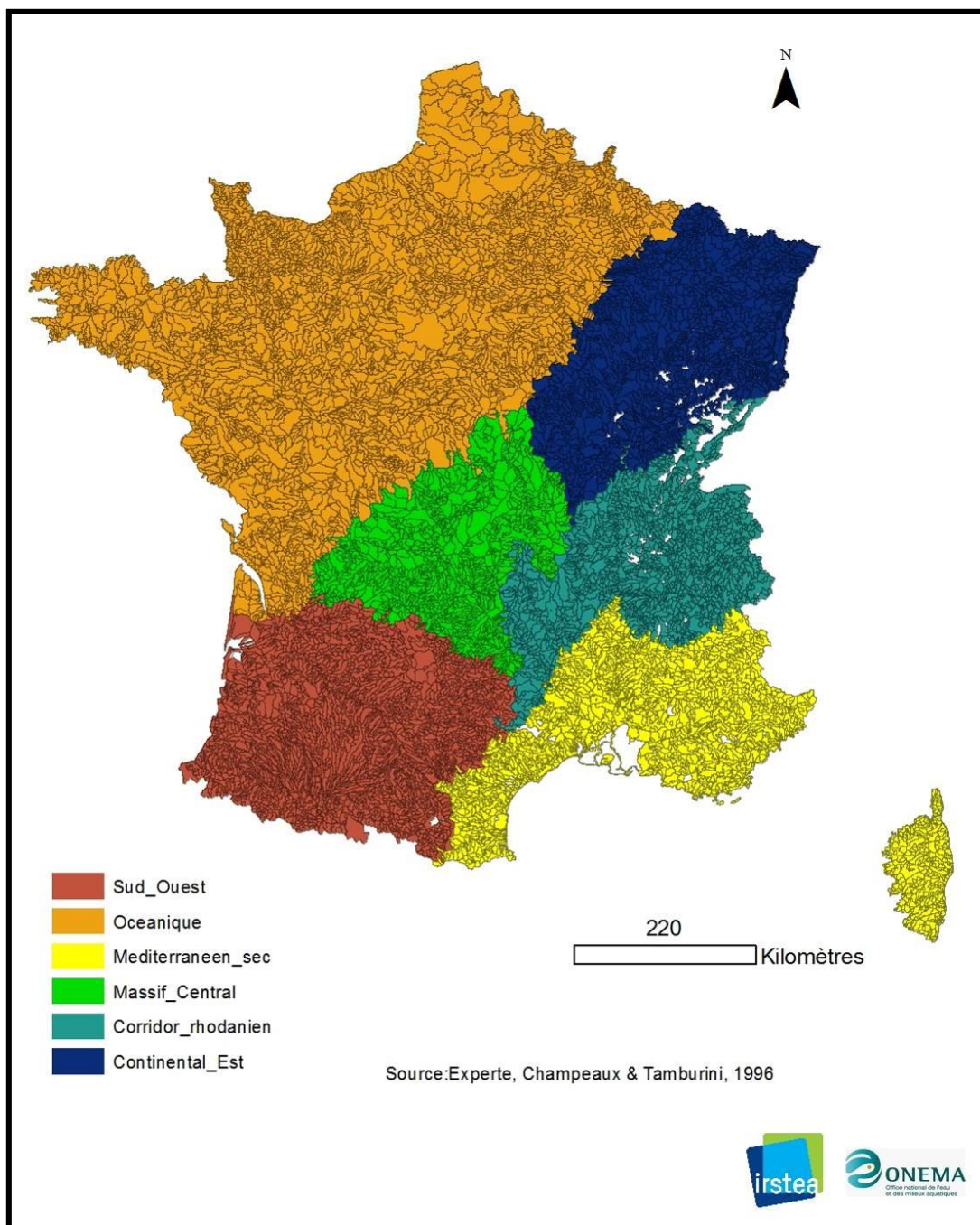
Désagrégation par jointure spatiale

9.10.4. Rappel des modalités que peut prendre la variable dans ARPEGES

Chacune des 6 zones correspond à une modalité.

9.10.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carte ci-dessous présente les zones hydrographiques pour chaque BVME.



Zone climatique considérée par le modèle ARPEGES par bassin versant local des masses d'eaux

9.11. Annexe 11 : Fiche synthétique pour estimer la variable « Cumul de pluies »

9.11.1. Objectifs

Caractériser le cumul de pluie mensuel moyen pour chacune des 2 périodes de l'année identifiées (novembre à mars et avril à octobre).

9.11.2. Données sources

Cumul mensuel des précipitations interpolées par mailles de 1 km² (méthode Aurhely de Météo France), normales 1971-2000.

9.11.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 1 : le BVME contient les entités de la donnée

Méthode :

Agrégation : Moyenne des mailles qui appartiennent au BVME.

Ensuite, la variable est normalisée pour que les valeurs prises soient comprises entre 0 et 100. Si on appelle cette variable X, on considère que X=100 correspond à un cumul de 300 mm et X=0 à un cumul nul (la valeur de 300 mm a été choisie car le cumul mensuel maximum en une maille tous mois confondus est de 302 mm). Le calcul est donc :

$$X = \text{Cumul mensuel} * 100 / 300$$

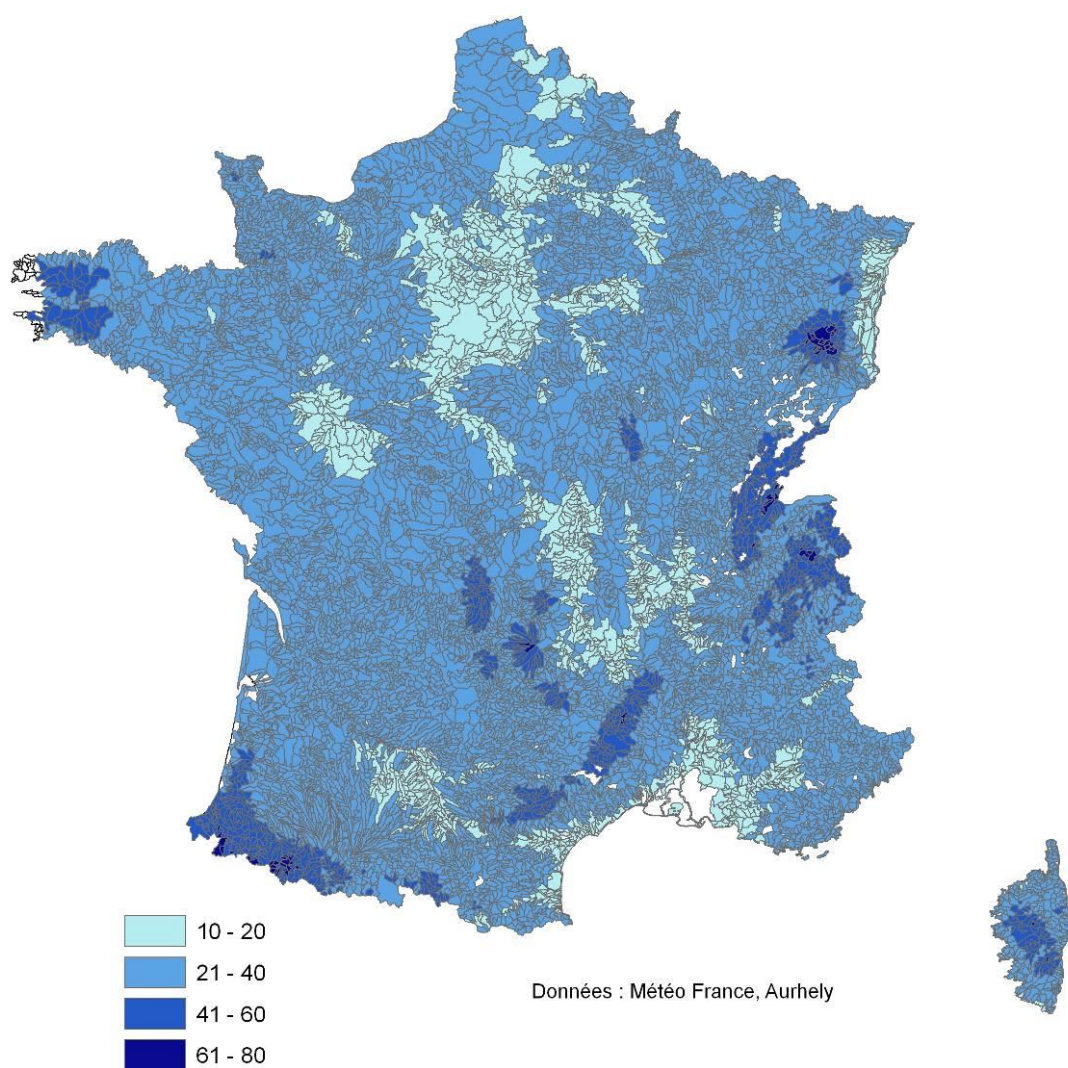
(X = 10 correspond donc à un cumul de 30 mm, X = 50 à un cumul de 150 mm, etc.)

Cette variable comprend deux modalités :

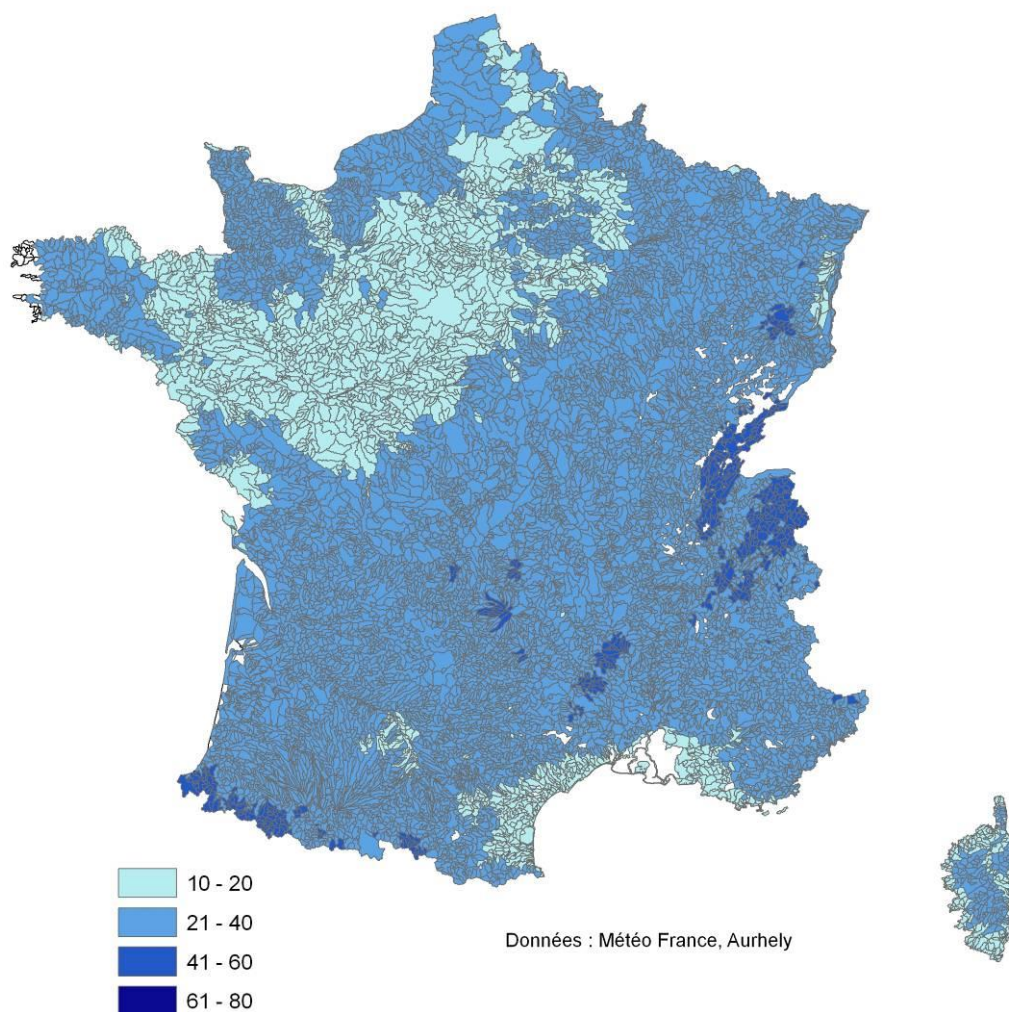
Cumul fort : X

Cumul faible : 100-X

9.11.4. Résultat de la discrétisation – Carte en entrée d'ARPEGES



Valeur de la modalité « cumul fort » de la variable « cumul de pluie » pour la période de novembre à mars



Valeur de la modalité « cumul fort » de la variable « cumul de pluie » pour la période d'avril à octobre

9.12. Annexe 12 : Fiche synthétique pour estimer la variable « Teneur en matière organique du sol »

9.12.1. Objectifs

Caractériser la teneur en matière organique dans les sols pour chaque BVME.

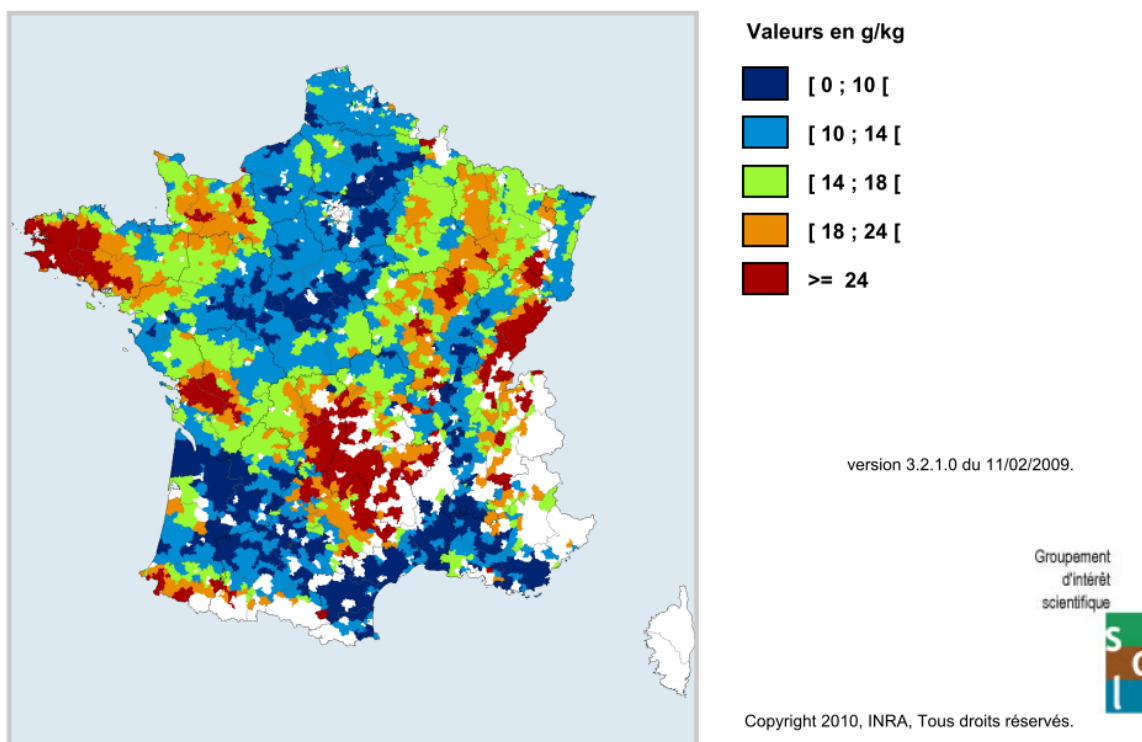
9.12.2. Données sources

Teneur en carbone organique, base de données des analyses de terre (BDAT) INRA 2010. La teneur en carbone organique est exprimée en g/kg.

On ne dispose pas de données pour certains cantons de montagne (pas assez d'analyses de terre réalisées). On a affecté à ces cantons la moyenne des autres cantons de la petite région agricole (PRA) à laquelle ils appartiennent. Si aucun canton de la PRA n'était renseigné, on leur a affecté la moyenne des cantons voisins. Une teneur faible a été affectée aux cantons corses (la plupart des cantons du pourtour méditerranéen ont une teneur faible).

La carte avant discrétisation est la suivante :

**Médiane des teneurs en carbone organique de l'horizon de surface des sols agricoles
(valeurs obtenues par oxydation humide) -- Période début 2000 à fin 2004**



9.12.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 2 : le BVME et les entités de la donnée se recouvrent

Méthode :

Moyenne des teneurs des cantons intersectant le BVME pondérées par la surface de ces intersections

9.12.4. Rappel des modalités que peut prendre la variable dans ARPEGES

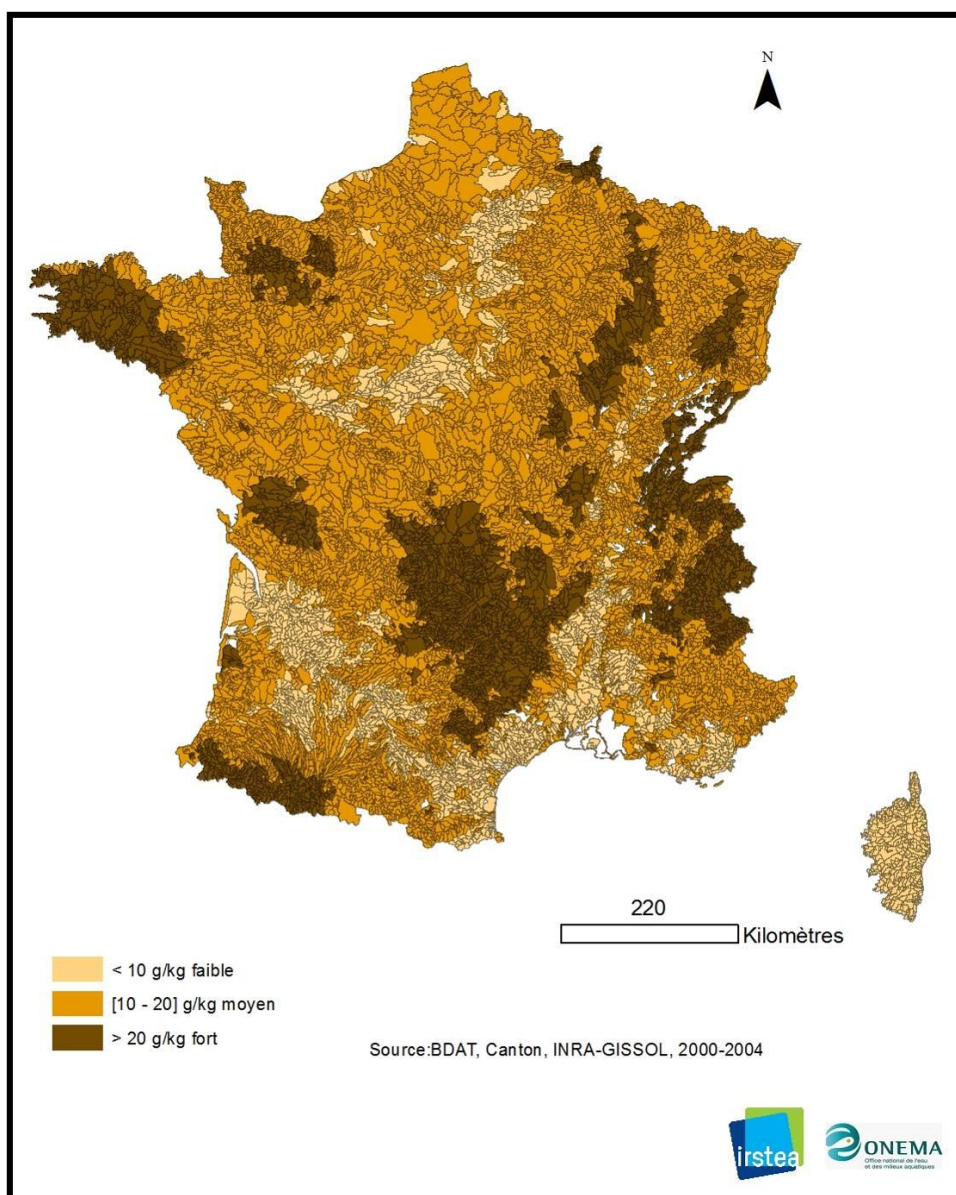
<10 g/kg → faible

Entre 10 et 20 → moyen

>= 20 → fort

9.12.1. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carte ci-dessous présente la teneur en carbone organique dans les sols à l'échelle du bassin versant local des masses d'eaux discrétisée selon les modalités d'entrée du modèle ARPEGES



Teneur en matière organique dans les sols à l'échelle du bassin versant local des masses d'eaux discrétisé selon les modalités d'entrée du modèle ARPEGES

9.13. Annexe 13 : Fiche synthétique pour estimer la variable « Pression d'usage phytosanitaire »

9.13.1. Objectifs

Caractériser les paramètres climatiques qui entrent en jeu dans les phénomènes de ruissellement, notamment l'intensité, la fréquence et le cumul des pluies.

9.13.2. Données sources

Les 6 zones climatiques ont été définies à l'aide d'un zonage existant basé sur les séries de précipitations (Champeaux & Tamburini 1996). Les zones ainsi obtenues à dire d'experts correspondent à des zones globalement homogènes du point de vue de la fréquence et de l'intensité des pluies.

9.13.3. Calcul à l'échelle du BVME

Cas de relation spatiale entre la donnée source et les BVME :

Cas n° 3 : le BVME est complètement contenu dans les entités de la donnée

Méthode :

Désagrégation par jointure spatiale

9.13.4. Rappel des modalités que peut prendre la variable dans ARPEGES

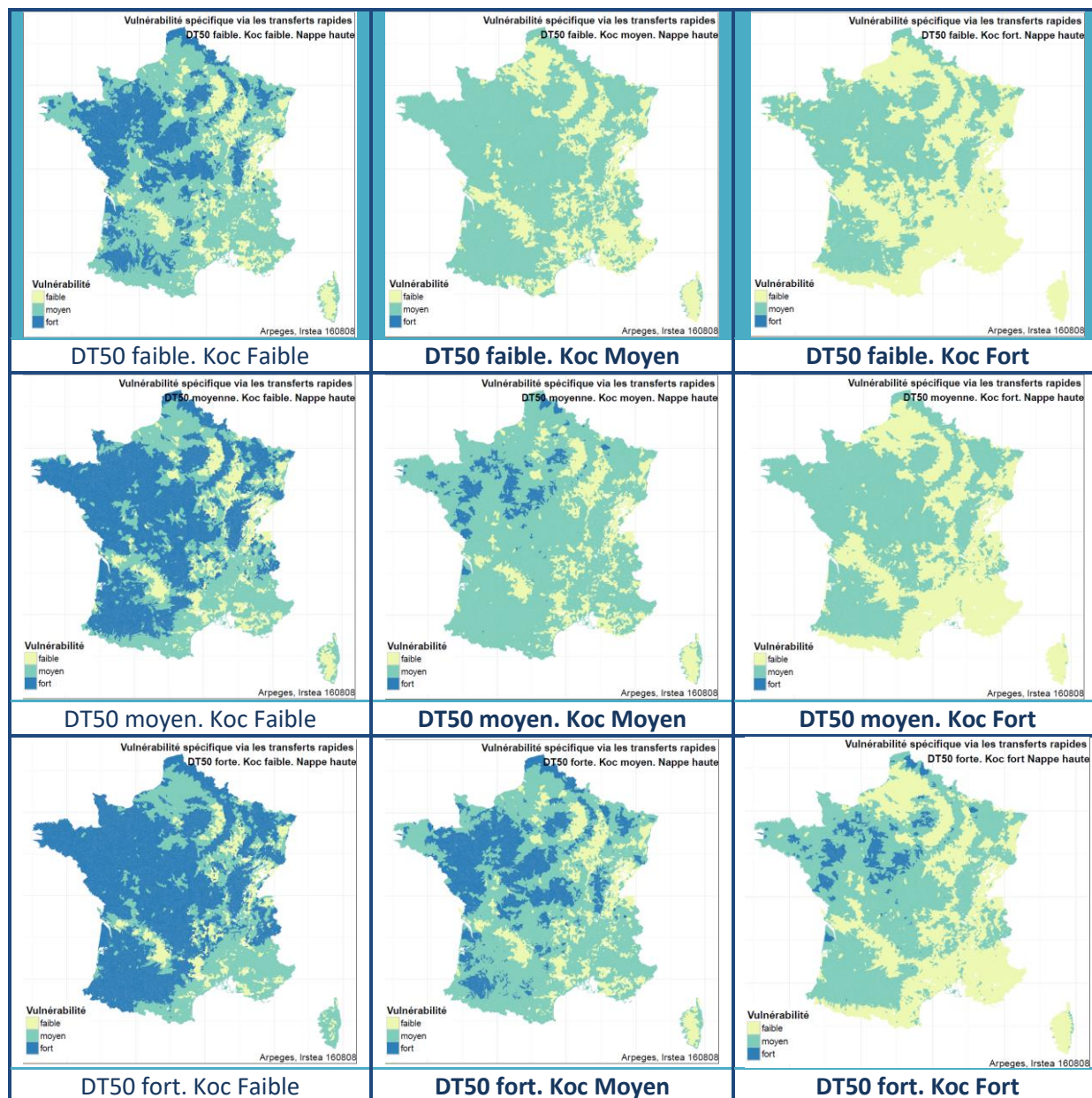
Chacune des 6 zones correspond à une modalité.

9.13.5. Résultat de la discrétisation – Carte en entrée d'ARPEGES

La carte ci-dessous présente les zones hydrographiques pour chaque BVME.

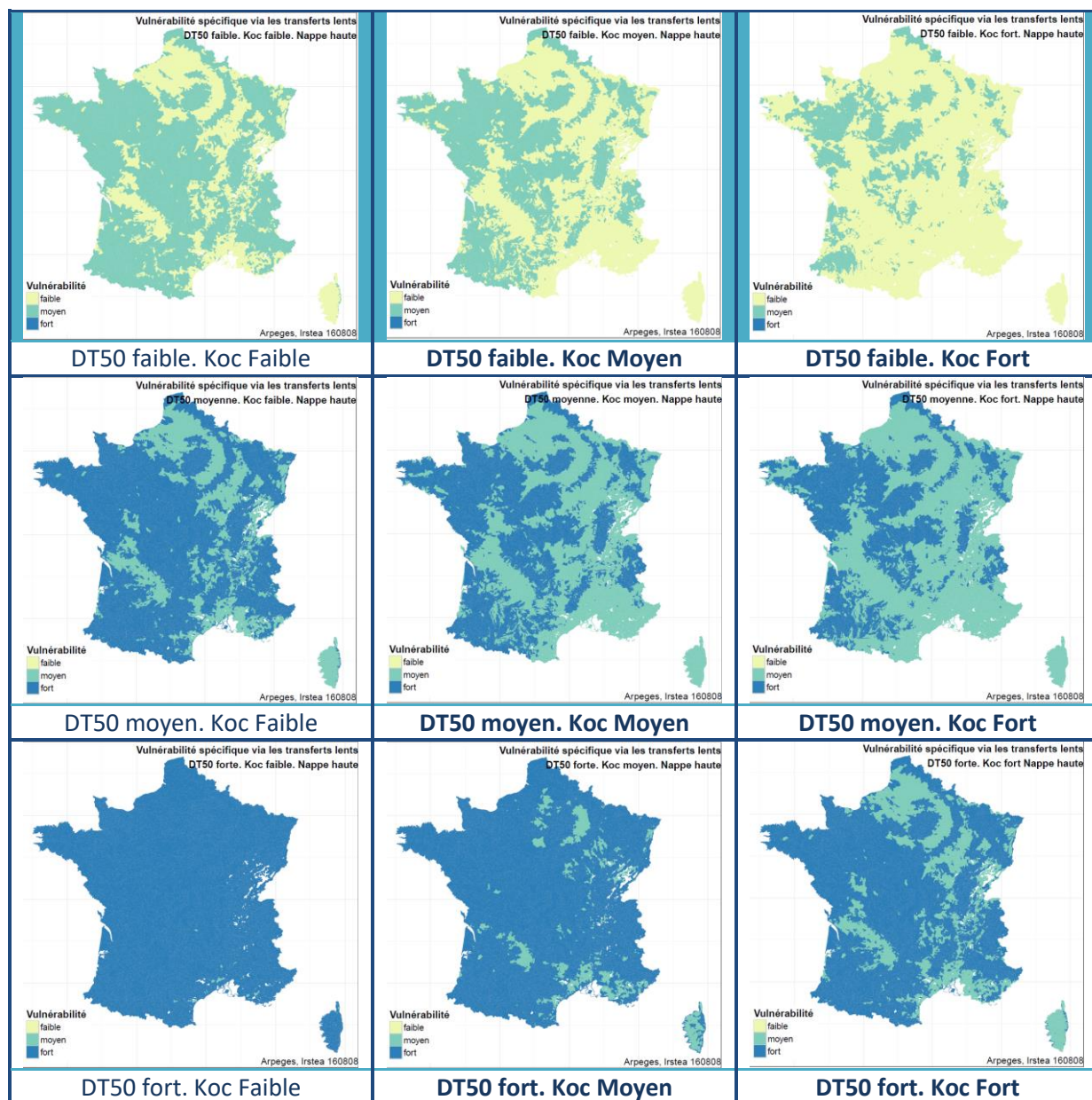
9.14. Annexe 14 : Vulnérabilités spécifiques pour les 9 combinaisons de caractéristiques physico-chimiques des substances

9.14.1. Période de Nappe Haute. Transferts rapides



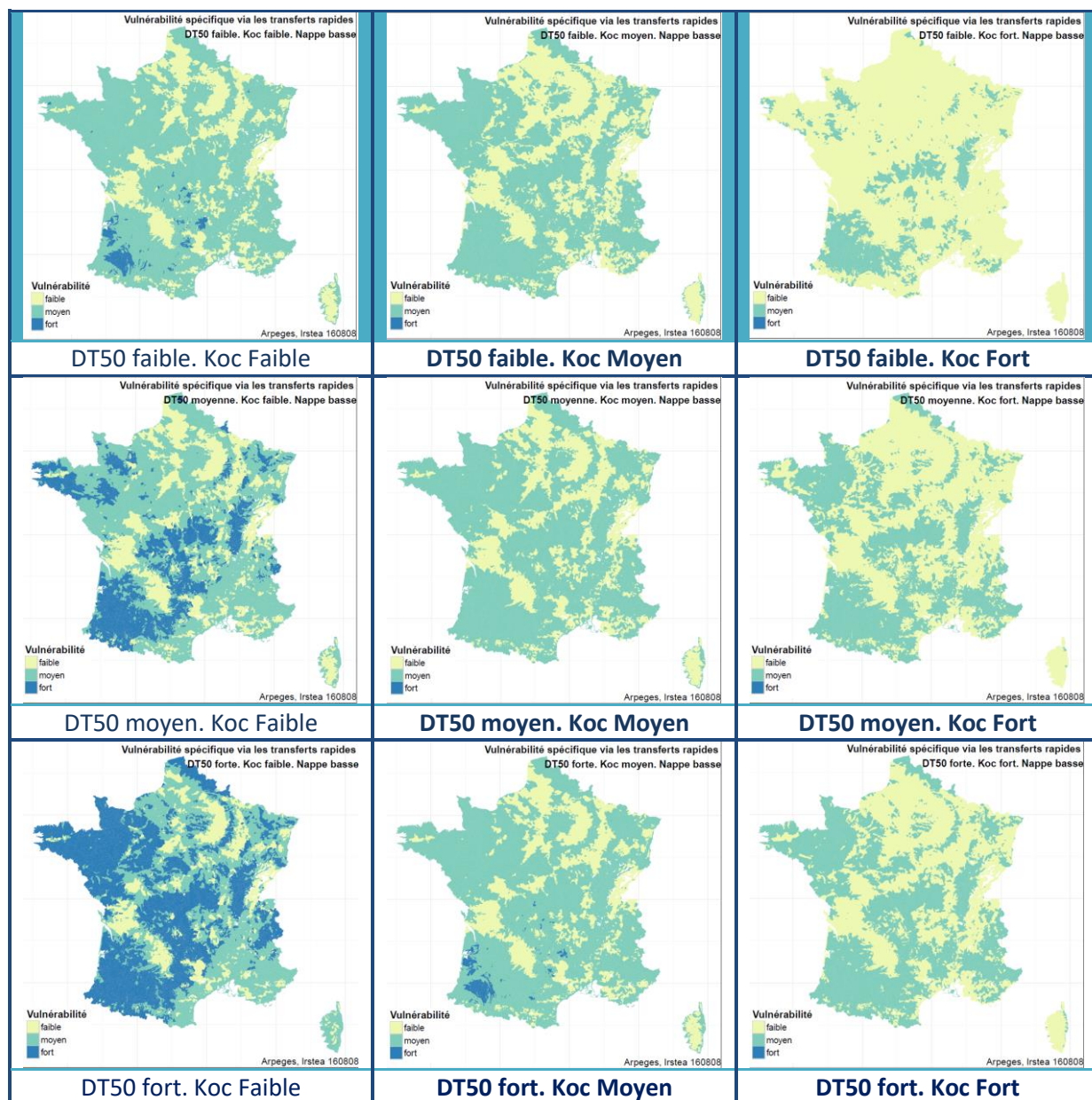
Vulnérabilité spécifique aux transferts rapides en période de Nappe Haute

9.14.2. *Période de Nappe Haute. Transferts lents*



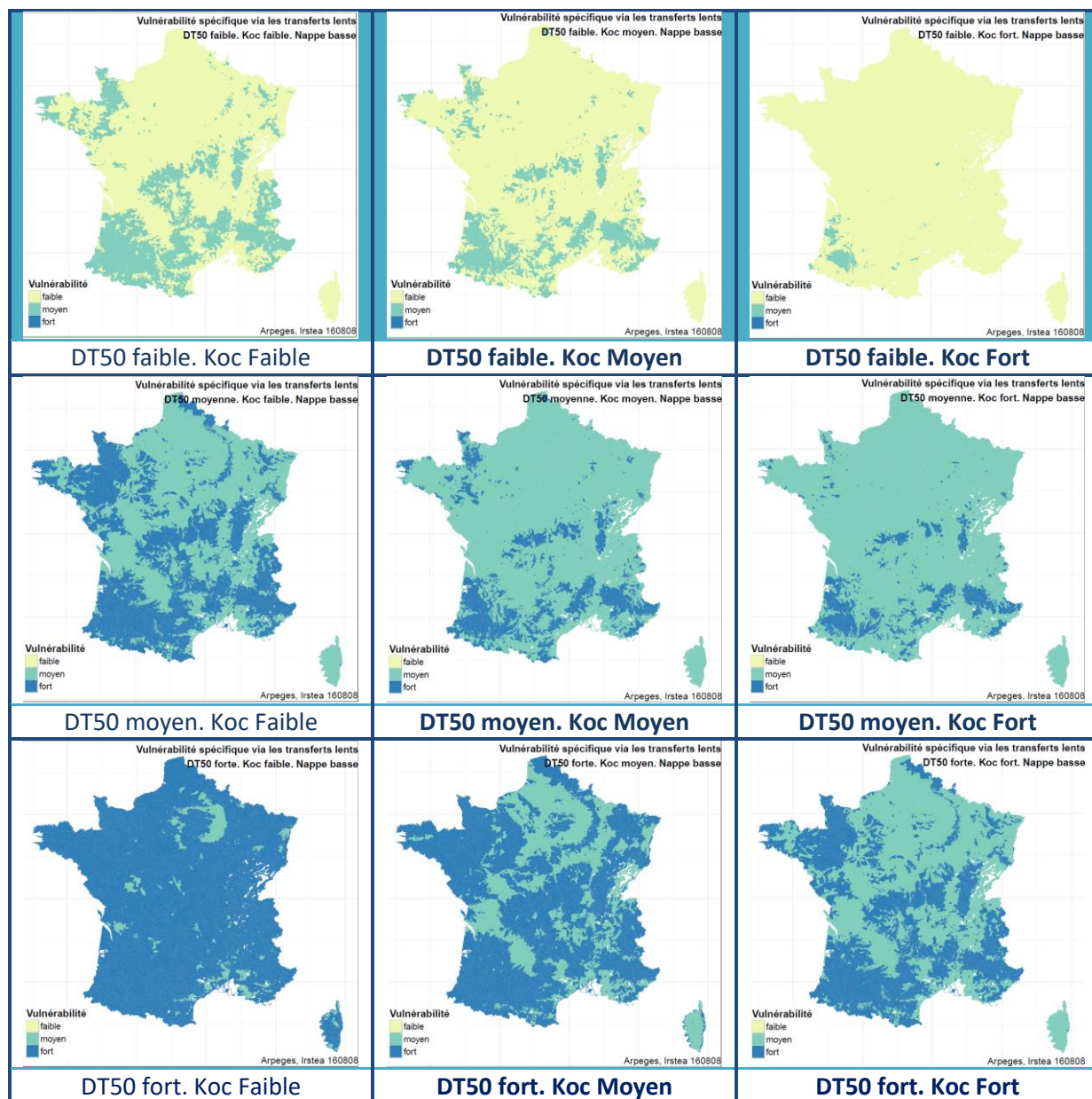
Vulnérabilité spécifique aux transferts lents en période de Nappe Haute

9.14.3. *Période de Nappe Basse. Transferts rapides*



Vulnérabilité spécifique aux transferts rapides en période de Nappe Basse

9.14.4. Période de Nappe Basse. Transferts lents



Vulnérabilité spécifique aux transferts lents en période de Nappe Basse